

N° d'ordre:

AD-A244 043



THESE

DAJ445-90-C-0008

présentée à

**L'UNIVERSITE DE PROVENCE
CENTRE SAINT-CHARLES**

**DTIC
ELECTE
DEC 7 1991
S C D**

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITE DE PROVENCE

Spécialité :

MATHEMATIQUES APPLIQUÉES

par

PAULA M. L. P. MILHEIRO de OLIVEIRA

Sujet de la thèse :

**ÉTUDES ASYMPTOTIQUES EN FILTRAGE NON LINEAIRE
AVEC PETIT BRUIT D'OBSERVATION**

Soutenue le 26 Septembre 1990 devant le jury composé de :

M. B. COUPET

Président

Mme D. MICHEL

Examineurs

MM. E. PARDOUX

J. PICARD

W. RUNGGALDIER

91-17370



91 1209 036

N° d'ordre:

THESE

présentée à

L'UNIVERSITE DE PROVENCE
CENTRE SAINT-CHARLES

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITE DE PROVENCE

Spécialité :

MATHEMATIQUES APPLIQUÉES

par

PAULA M. L. P. MILHEIRO de OLIVEIRA

Sujet de la thèse :

*Asymptotic Studies in Nonlinear Time Filtering
with small observation Noise)*

ÉTUDES ASYMPTOTIQUES EN FILTRAGE NON LINEAIRE
AVEC PETIT BRUIT D'OBSERVATION

Soutenue le 26 Septembre 1990 devant le jury composé de :

M. B. COUPET

Président

Mme D. MICHEL

Examineurs

MM. E. PARDOUX

J. PICARD

W. RUNGALDIER

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By <i>Per Form 50</i>	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
<i>A-1</i>	



Je tiens à remercier Etienne Pardoux, Professeur à l'Université de Provence, d'avoir accepté de diriger ma thèse, de m'avoir proposé des problèmes si intéressants et de m'avoir encouragée dans mon travail. Je dois remercier de la même façon Jean Picard, Chercheur à l'INRIA, pour tout le soin porté à orienter ces travaux et pour toutes les idées apportées dans leur réalisation. Je le remercie de sa disponibilité et pour le temps consacré à discuter ce sujet, à lire et à corriger une grande partie de cette thèse.

Je remercie également les membres du jury et les rapporteurs d'avoir accepté la tâche qui leur a été confiée.

Je remercie l'équipe MEFISTO de l'ambiance sympathique qui régna tout au long de ces trois ans. Spécialement, je dois exprimer ma gratitude à Fabien Campillo et François Le Gland pour leur disponibilité et pour l'aide apportée à plusieurs reprises.

A l'ensemble du personnel INRIA je dois un remerciement pour leur accueil.

Je remercie chaleureusement Ricardo Lima pour toute son aide. Son soutien et ses conseils ont été d'une grande importance au moment le plus difficile de mon adaptation en France et du choix de mes études.

Je veux aussi exprimer ma reconnaissance à l'Ecole d'Ingénieurs de l'Université de Porto, au Département de Génie Civil, en la personne de Monsieur Tentugal Valente, pour la manière enthousiaste avec laquelle il a soutenu ce projet.

Je ne dois pas oublier de remercier le Groupe de Mathématiques Appliquées de l'Université de Porto de m'avoir donné l'occasion de préparer ma thèse en France.

Finalement, je garde un remerciement très spécial à la personne qui a suivi de tout près mes premiers pas dans les Probabilités. A Francisco Calheiros, Professeur à l'Ecole d'Ingénieurs de l'Université de Porto, je veux exprimer ma profonde reconnaissance de m'avoir ouvert le chemin dans mon "odyssée", de m'avoir toujours encouragée au cours de ces quatre années d'études et pour ses conseils si précieux.

Je ne pourrais pas oublier mes parents, ma sœur et tous les amis qui ont fait que le temps se soit écoulé si vite.

Introduction

Nous considérons le problème de filtrage non linéaire dans une situation particulière, celle de l'observation faiblement bruitée. En pratique, plusieurs problèmes de filtrage, en particulier en génie aérospatial, peuvent être modélisés de façon convenable par des situations de ce type.

Ce travail est consacré à l'étude de problèmes asymptotiques. Nous considérons des filtres sous-optimaux, dépendant d'un paramètre ϵ et nous étudions leur comportement lorsque ϵ converge vers zéro.

Il est bien connu que le problème de filtrage non linéaire est un problème de dimension infinie, sa solution étant donnée par exemple par l'équation de Zakai. Quelques tentatives ont été faites pour trouver des classes de problèmes pour lesquels un filtre de dimension finie existe (cf. Benès [1], Levine [19] et Wong [30], entre autres). Cependant ces tentatives se sont montrées d'un intérêt limité, n'ayant pas abouti à des classes de problèmes "intéressants". Des travaux postérieurs viendraient prouver qu'on ne peut construire un filtre de dimension finie que pour une classe de problèmes négligeable. Les conditions que ces problèmes doivent vérifier sont liées à l'existence d'une algèbre de Lie de dimension finie (cf. par exemple Chaleyat-Michel [5]). On regarde donc avec intérêt la possibilité d'obtenir de "bonnes" approximations du filtre optimal qui puissent être calculées à l'aide d'un nombre fini de statistiques données comme des solutions d'un système d'équations différentielles stochastiques conduites par le processus d'observation.

De nombreux travaux ont été effectués dans ce domaine - voir par exemple Fleming-Ji-Pardoux [9], Fleming-Ji-Salame-Zhang [10] et Roubaud [28] pour le problème linéaire par morceaux, Fleming-Pardoux [11] pour le problème monotone par morceaux, Katzur-Bobrovsky-Schuss [17], Picard [22] et [23] et Bensoussan [2] pour le problème asymptotiquement observé, Yaesh-Bobrovsky-Schuss [31] et Picard [26] pour le problème asymptotiquement non observé, entre autres.

Cette thèse peut être vue comme composée de trois chapitres indépendants. Cependant tous les trois s'inscrivent dans un même sujet: le comportement asymptotique de problèmes de filtrage.

Dans un premier chapitre nous étudions un problème de filtrage non linéaire unidimensionnel en temps discret, lorsque le bruit d'observation est faible. Il se

caractérise par une fonction d'observation injective.

Le deuxième chapitre est consacré à un problème de filtrage linéaire par morceaux en temps discret. Nous nous intéressons aux aspects de mise en œuvre d'algorithmes pour résoudre numériquement ce problème. Ce travail a été fait en collaboration avec M.C. Roubaud.

Dans le dernier chapitre nous étudions un problème en temps continu, ayant des dimensions différentes pour le signal et l'observation. Plus exactement nous considérons un signal de dimension 2 dont seulement une des composantes est bruitée et un processus d'observation de dimension 1 observant l'autre composante.

Filtres approchés pour un problème de filtrage non linéaire en temps discret avec petit bruit d'observation

- [1] *Filtres approchés pour un problème de filtrage non linéaire en temps discret avec petit bruit d'observation*, RR 1142, INRIA, Dec. 1989.
- [2] *Filtres approchés pour un problème de filtrage non linéaire en temps discret avec petit bruit d'observation*, Lecture N. in Cont. and Inf. Sci., Proc. of the 9th Internat. Conf. on Analysis and Optim. of Systems, Antibes, June 12-15, 1990, pp. 198-207.

On considère le système différentiel stochastique

$$\begin{cases} X_{k+1} = X_k + b(X_k) \Delta t + \sigma \sqrt{\Delta t} w_{k+1}, \\ y_k = h(X_k) + \frac{\varepsilon}{\sqrt{\Delta t}} \bar{w}_k, \end{cases}$$

où $\{w_k\}$ et $\{\bar{w}_k\}$ sont des bruits blancs gaussiens indépendants, ε est un paramètre supposé "petit" et Δt dépend de ε . Le processus $\{X_k\}$ est le signal à estimer et $\{y_k\}$ est le processus d'observation engendrant la filtration \mathcal{Y}_k .

On suppose que la fonction h est injective, i.e. le signal est asymptotiquement observé. On cherche des approximations M_k de

$$\hat{X}_k = \mathbb{E}[X_k | \mathcal{Y}_k]$$

qui soient données de façon récursive et qui puissent être calculées, à l'instant k , à partir de l'observation y_k . Pour ces processus $\{M_k\}$, on veut pouvoir estimer $\hat{X}_k - M_k$ lorsque ε converge vers 0.

Nous commençons par l'étude du cas linéaire et nous y trouvons l'inspiration nécessaire pour construire les filtres approchés dans le cas non linéaire. Nous exploitons l'idée de Picard [22] et [23] pour le problème en temps continu et nous l'adaptions au problème en temps discret. Nous faisons appel à des outils comme les changements de probabilité et une version discrète du Théorème de Girsanov,

la dérivation par rapport à la condition initiale et une formule d'intégration par parties. Nous démontrons que, si $\Delta t = \varepsilon^\alpha$ avec $\alpha > 1$, le filtre

$$M_{k+1} = M_k + b(M_k) \Delta t + \sigma \frac{\Delta t}{\varepsilon} (\bar{y}_{k+1} - h(M_k)),$$

avec condition initiale $M_0 = m_0$ quelconque, approche \hat{X}_k à l'ordre $\mathcal{O}(\varepsilon \vee \varepsilon^{\alpha-1/2})$; si $\alpha = 1$, le filtre de dimension 2

$$\begin{aligned} M_{k+1} &= M_k + b(M_k) \varepsilon + \bar{\theta}_k (\bar{y}_{k+1} - h(M_k)) \\ \bar{\theta}_k &= \frac{h'(M_k) [\bar{\theta}_{k-1} + \sigma^2 h'(M_{k-1})]}{h''(M_k) [\bar{\theta}_{k-1} + \sigma^2 h'(M_{k-1})] + h'(M_{k-1})}, \end{aligned}$$

avec condition initiale $M_0 = m_0$, $\bar{\theta}_0 = \theta_0$ quelconque telle que $\sigma \theta_0 > 0$, approche le filtre optimal à l'ordre $\mathcal{O}(\varepsilon)$; si $\alpha < 1$, il est suffisant d'inverser la fonction d'observation.

Les estimations obtenues vérifient la propriété d'uniformité par rapport au temps et montrent que ces filtres approchés ont peu de mémoire.

Nous étudions ensuite un cas particulier où la fonction d'observation est supposée non injective. Il s'agit du cas d'un système linéaire par morceaux.

Filtrage linéaire par morceaux d'un système en temps discret avec petit bruit d'observation

Nous considérons le système linéaire par morceaux suivant:

$$\begin{cases} X_{k+1} = x_k + \varepsilon b(X_k) + \sqrt{\varepsilon} \sigma(X_k) u_k \\ y_k = h(X_k) + \sqrt{\varepsilon} v_k, \end{cases}$$

où $\{w_k\}$ et $\{v_k\}$ sont des bruits blancs gaussiens indépendants et

$$\begin{aligned} b(x) &= B_- x \mathbf{1}_{\{x < 0\}} + B_+ x \mathbf{1}_{\{x \geq 0\}} \\ \sigma(x) &= \sigma_- \mathbf{1}_{\{x < 0\}} + \sigma_+ \mathbf{1}_{\{x \geq 0\}} \\ h(x) &= H_- x \mathbf{1}_{\{x < 0\}} + H_+ x \mathbf{1}_{\{x \geq 0\}}. \end{aligned}$$

Si $H_+ H_- > 0$, i.e. la fonction h est monotone, le problème n'offre aucune difficulté. Nous nous intéressons au cas $H_+ H_- < 0$, i.e. h n'est pas monotone. Dans ce cas, bien que $h(X_t)$ puisse être facilement estimé il n'en est pas de même pour X_t . L'idée de base est la suivante: dès que le signal reste dans une région où la fonction d'observation est injective (voir linéaire), on peut appliquer un filtre classique, type Kalman-Bucy. Le problème se réduit donc à la recherche d'une méthode pour

détecter les changements de signe du processus $\{X_k\}$, ainsi que déterminer le signe pris dans chacun des intervalles ainsi séparés.

Cette idée avait déjà été exploitée dans Fleming-Ji-Pardoux [9], pour un système linéaire par morceaux, sous une certaine hypothèse de "déteclabilité" et, plus tard, par Fleming-Pardoux [11] pour un système non linéaire par morceaux. Le problème en temps discret a été abordé dans Fleming-Ji-Salame-Zhang [10]. Nous étudions ici un problème plus général que celui traité par ces auteurs. En effet, nous considérons une hypothèse plus large que l'hypothèse de "déteclabilité" émise dans l'étude précédente. Nous supposons que

$$\begin{cases} H_-^2 \sigma_-^2 \neq H_+^2 \sigma_+^2 & (HD1) \\ \text{ou} \\ H_-^2 \sigma_-^2 = H_+^2 \sigma_+^2 \text{ et } B_+ \neq B_- & (HD2). \end{cases}$$

Ceci permet de traiter en particulier le cas d'une fonction d'observation symétrique par rapport à l'axe des ordonnées.

Les aspects de mise en œuvre d'algorithmes pour résoudre ce problème sont abordés.

Nous présentons des tests permettant de séparer les intervalles où le signal est positif ou négatif et de déterminer ce signe. Un premier type de tests détecte les changements de signe de $\{X_k\}$. Deux tests sont proposés: un test sur les observations et un test sur les sorties des filtres de Kalman-Bucy. Le premier de ces tests a déjà été étudié dans [10], sous l'hypothèse (HD1). Le deuxième type de tests vise la détermination du signe de $\{X_k\}$ dans l'intervalle de monotonie détecté par un test du premier type. Sous l'hypothèse (HD1), deux tests ont été décrits dans [10]: un test de variation quadratique et un test de rapport de vraisemblance sur les accroissements des observations. Sous l'hypothèse (HD2), cependant, le premier de ces tests ne permet pas de prendre une décision et on se contente donc d'adapter le deuxième. Nous décrivons un nouveau test de rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy (cf. Roubaud [28], pour le problème en temps continu).

L'application de ces tests repose sur une méthode efficace de détermination des bornes. Nous présentons une étude purement heuristique permettant d'obtenir des formules pour ces quantités, ainsi que pour les temps d'attente espérés dans les tests de décision sur le signe.

Nous comparons les différents tests et étudions l'influence des paramètres du système sur les résultats obtenus sur des exemples. En particulier nous remarquons que les temps d'attente d'un test sous l'hypothèse (HD2) sont beaucoup plus longs que ceux sous l'hypothèse (HD1).

Nous étudions ensuite une autre situation où la fonction d'observation n'est pas injective. Il s'agit d'un cas où le signal est un processus bidimensionnel et

l'observation est un processus unidimensionnel.

Filtres approchés pour un problème de filtrage non linéaire d'une diffusion de dimension 2 mesurée par un processus de dimension 1 faiblement bruité

Nous nous intéressons au problème de filtrage associé au système différentiel stochastique, au sens d'Itô

$$\begin{cases} dx_1(t) = f_1(x_1(t), x_2(t)) dt \\ dx_2(t) = f_2(x_1(t), x_2(t)) dt + \sigma(x_1(t), x_2(t)) dw(t), \\ dy(t) = h(x_1(t)) dt + \varepsilon d\bar{w}(t), \end{cases}$$

Nous cherchons donc à estimer le signal $\{X_t = (x_1(t), x_2(t))\}$, à valeurs dans \mathbb{R}^2 , à l'aide du processus d'observation y_t , processus à valeurs dans \mathbb{R} . Notre but est d'étudier le comportement asymptotique d'un tel estimateur, lorsque ε tend vers 0. Comme dans Picard [26], nous privilégions des méthodes permettant de déterminer l'ordre de grandeur de l'erreur de filtrage, voir sa vitesse de convergence. Nous y trouvons l'inspiration nécessaire.

Nous traitons d'abord le cas où la fonction h est linéaire et f_1 est linéaire par rapport à x_2 . Nous considérons un filtre approché $\{M_k\}$ du type proposé dans Yaesh-Brobovsky-Schuss [31]:

$$\begin{cases} dm_1(t) = f_1(m_1(t), m_2(t)) dt + \sqrt{\frac{2\bar{\sigma}F_{12}}{\varepsilon H}} [dy(t) - Hm_1(t) dt] \\ dm_2(t) = f_2(x_1(t), x_2(t)) dt + \frac{\bar{\sigma}}{\varepsilon} [dy(t) - Hm_1(t) dt], \end{cases}$$

avec condition initiale M_0 et où $\bar{\sigma}$ est une constante strictement positive. Nous montrons que le moment d'ordre $2p$ de $X_k - M_k$ est asymptotiquement d'ordre $\mathcal{O}(\varepsilon^{p/2})$, pour la première composante, et d'ordre $\mathcal{O}(\varepsilon^{p/2})$ pour la seconde. Nous démontrons que la condition initiale disparaît exponentiellement vite.

Nous traitons ensuite le problème général. Par une transformation de coordonnées nous nous ramenons au cas étudié précédemment. Nous obtenons les mêmes résultats pour la covariance de $X_k - M_k$.

Sommaire

- 1. Filtres approchés pour un problème de filtrage non linéaire en temps discret avec petit bruit d'observation**
- 2. Filtrage linéaire par morceaux d'un système en temps discret avec petit bruit d'observation**
- 3. Filtres approchés pour un problème de filtrage non linéaire d'une diffusion de dimension 2 mesurée par un processus de dimension 1 faiblement bruité**

1

**Filtres approchés pour un
problème de filtrage non linéaire
en temps discret avec petit bruit
d'observation**

FILTRES APPROCHÉS POUR UN PROBLÈME DE FILTRAGE NON LINÉAIRE DISCRET AVEC PETIT BRUIT D'OBSERVATION

APROXIMATE FILTERS FOR A NONLINEAR DISCRETE TIME FILTERING PROBLEM WITH SMALL OBSERVATION NOISE

Résumé

On fait une étude asymptotique d'un problème de filtrage non linéaire en temps discret unidimensionnel, lorsqu'un certain paramètre ε tend vers 0. On traite le cas d'un problème de filtrage non linéaire en temps discret issu d'un problème en temps continu avec petit bruit d'observation. Des filtres approchés de dimension finie sont proposés et les estimations sur les performances de ces filtres sont énoncées et démontrées. On commence par faire une étude dans le cas linéaire, ce qui va nous inspirer les filtres approchés à prendre dans le cas général. Pour le résultat concernant l'erreur entre le filtre approché et le filtre optimal, la démonstration présentée utilisera des changements de probabilité et la dérivation par rapport à la condition initiale. Pour terminer, on présente des résultats d'application de ces filtres à un exemple et on constate que les propositions énoncées sont vérifiées numériquement.

Abstract

We study the asymptotic behaviour of a nonlinear one-dimensional filtering discrete time problem, as some parameter ε tends to 0. We treat the case of a nonlinear discrete time problem coming from a continuous time one with small observation noise. Finite dimensional approximate filters are proposed and results concerning estimations of their performance are stated and proved. We present, at first, a brief study of the linear case and get the necessary inspiration to construct the approximate filters in the general case. The proof of the result we give concerning the error between the approximate filter and the optimal one will make use of probability changes and differentiation with respect to the initial condition. Finally, we present the results obtained when applying those filters to an example and we notice that the propositions previously stated are verified numerically.

1 Introduction

On considère le problème de filtrage non linéaire en temps continu suivant :

On a un signal X , solution de l'équation différentielle stochastique unidimensionnelle

$$dX_t = b(X_t) dt + c dw_t^1, \quad X_0 = \xi \quad (1)$$

et on dispose de l'observation unidimensionnelle vérifiant

$$dY_t = h(X_t) dt + \epsilon dw_t^2, \quad Y_0 = 0 \quad (2)$$

où w^1, w^2 sont des processus de Wiener standards indépendants et ξ est une variable aléatoire indépendante de w^1 et w^2 . Le paramètre ϵ est supposé petit.

Le problème de filtrage consiste à calculer, pour toute fonction mesurable ϕ , l'espérance conditionnelle de $\phi(X_t)$, étant donnée l'observation Y , jusqu'à l'instant t .

On aimerait résoudre numériquement ce problème. Une possibilité est donnée par la résolution de l'équation de Zakai, une fois discrétisée, ce qui a été étudié dans [Le Gland]. Cependant un tel calcul présente une grande complexité, puisqu'il s'agit, dans le cas général, d'une équation de dimension infinie. Une autre possibilité consiste à utiliser les filtres approchés proposés dans [Picard] et les discrétiser en temps. Ces filtres étant de dimension finie, le problème devient plus simple à résoudre. De plus, on y trouve des résultats concernant l'erreur de filtrage et en particulier ceux qui correspondent au filtre de Kalman étendu. La méthode de démonstration utilisée fait appel à des changements de probabilité, à la théorie de retournement du temps pour les diffusions et à des résultats sur les flots définis par des équations différentielles stochastiques. Dans [KBS] on peut trouver les premiers résultats sur ce genre de filtres, cependant leur justification reste purement formelle. Notons également que dans [Bensoussan], l'auteur considère le même problème de filtrage en temps continu et utilise une méthode autre que celle développée dans [Picard] pour démontrer les résultats asymptotiques.

En pratique, néanmoins, on ne dispose souvent que de l'observation à des instants discrets et, de plus, pour mener des calculs sur ordinateur, il est nécessaire de manipuler des processus discrets. Au lieu de discrétiser le filtre on peut penser à discrétiser le problème lui même. L'avantage de cette approche réside dans le fait que, dès qu'on se limite à des processus en temps discret, on n'a plus besoin de la théorie des flots ou du retournement du temps et même pas de l'intégrale d'Itô.

Cette idée a motivé l'étude qui suit, dans laquelle on recherche des filtres approchés pour un problème de filtrage non linéaire discret avec petit bruit

d'observation.

On considère la discrétisation la plus simple des équations (1) et (2):

$$X_{k+1} = X_k + b(X_k) \Delta t + \sigma \sqrt{\Delta t} w_{k+1}, \quad X_0 = \xi \quad (3)$$

$$y_k = h(X_k) + \frac{\varepsilon}{\sqrt{\Delta t}} \bar{w}_k \quad (4)$$

où X_k est une approximation de X_{t_k} ($t_k = k \Delta t$), $\{w_k\}$ et $\{\bar{w}_k\}$ sont des bruits blancs gaussiens standards indépendants et ξ une v.a. indépendante de $\{w_k\}$ et $\{\bar{w}_k\}$. On se propose de résoudre, de façon approchée, le problème de filtrage associé au système discret (3 - 4), où Δt est pris comme étant un paramètre dépendant de ε .

Soit $\{\hat{X}_k\}$ le filtre optimal pour le problème discret, i.e.,

$$\hat{X}_k = E[X_k | Y^k], \quad Y^k = \sigma(y_i; i = 0, 1, \dots, k).$$

Puisque, dans le cas général, la détermination de \hat{X}_k présente une grande complexité, notre but sera de construire une "bonne approximation" de ce filtre représentant un compromis entre les coûts en temps de calcul et la précision des résultats. Si $\varepsilon = 0$ et h est inversible, le signal est observé exactement: $X_k = h^{-1}(y_k)$. Pour Δt fixé, quand $\varepsilon \rightarrow 0$, l'observation est faiblement bruitée et un "bon" filtre consistera à prendre tout simplement $M_k = h^{-1}(y_k)$. Cependant, notre intérêt portera sur des résultats asymptotiques: on fera tendre Δt vers 0 en même temps que ε et on donnera une approximation $\{M_k\}$ du filtre optimal en temps discret. On prendra $\Delta t = \varepsilon^\alpha$ et on s'intéressera à l'estimation de $|\hat{X}_k - M_k|$ quand ε devient "petit". Quand $\varepsilon \rightarrow 0$, on aura à distinguer 3 cas, selon la valeur de α . Les cas $\alpha > 1$, $\alpha < 1$ et le cas intermédiaire $\alpha = 1$ présentent un comportement différent. Quand α est "grand" on doit pouvoir approcher le problème de filtrage en temps continu (1 - 2). En revanche, quand α est "petit", le système (3 - 4) ne sera plus nécessairement une "bonne discrétisation" du système en temps continu donc le filtre n'aura pas le même comportement.

On commencera par étudier le cas linéaire dans la section 2, on proposera des filtres approchés unidimensionnels possibles et des résultats asymptotiques sur leur "qualité". Cette étude nous suggèrera le filtre approché à considérer dans le cas non linéaire, ce qui est fait dans la section 3, et la généralisation des résultats asymptotiques de "qualité" du filtre est obtenue.

On s'intéressera en premier à l'estimation de $X_k - M_k$, ce qui nous ramène simplement à l'étude d'une série récurrente. Ensuite, on procédera à l'estimation de $\hat{X}_k - M_k$. Dans le cas linéaire, puisqu'une équation pour \hat{X}_k est disponible (le filtre de Kalman-Bucy), l'étude d'une série récurrente permet encore d'aboutir au résultat voulu. Par contre, dans le cas non linéaire, la méthode sera toute

autre: on commencera par introduire des changements de probabilités, on utilisera une version discrète du Théorème de Girsanov (voir annexe A) et, par dérivation par rapport à la condition initiale, on se ramènera à l'étude d'une expression asymptotique pour $X_k - M_k$ dont on prendra l'espérance conditionnelle.

La section 3.4 contient un bref résumé sur les filtres approchés à considérer dans le cas non linéaire, selon la valeur de α , et des résultats sur leur efficacité.

On termine cette étude par un essai numérique, présenté dans la section 4.

Notation 1.1 *Dans la suite, on notera c ou C n'importe quelle constante positive indépendante de Δt et de ε .*

Remerciements: J'aimerais remercier E. Pardoux, Professeur à l'Université de Provence, pour m'avoir suggéré ce travail et J. Picard, chargé de recherche à l'INRIA, pour les suggestions et discussions fructueuses qui m'ont permis d'avancer dans ce sujet.

2 Cas linéaire

On considère le système linéaire:

$$\begin{cases} X_{k+1} = (1 + b \Delta t) X_k + \sigma \sqrt{\Delta t} w_{k+1}, & X_0 = \xi \\ y_k = h X_k + \frac{\varepsilon}{\sqrt{\Delta t}} \bar{w}_k \end{cases} \quad (5)$$

où

(H1) $\{w_k\}$ et $\{\bar{w}_k\}$ sont des bruits blancs gaussiens standards indépendants,

(H2) ξ est une v.a. gaussienne indépendante de $\{w_k\}$ et $\{\bar{w}_k\}$, de variance p_0 ,
où p_0 est une constante ne dépendant pas de ε ou de Δt .

(H3) $\sigma h > 0$.

Remarque 2.1 L'hypothèse (H3) peut, en fait, être remplacée par $\sigma h \neq 0$ et, dans ce cas, on remplacera, dans les expressions du gain, σh et σ/h par leurs valeurs absolues. Néanmoins l'hypothèse (H3) n'est pas plus restrictive que cette dernière, puisque si on remplace σ par $-\sigma$ dans le système (5) on obtient des modèles équivalents.

Il est bien connu que, dans le cas linéaire, la loi conditionnelle de X_k sachant Y^k est gaussienne et l'estimation optimale est donc donnée par des équations en dimension finie, les équations du filtre de Kalman-Bucy:

$$\hat{X}_k = (1 + b \Delta t) \hat{X}_{k-1} + \frac{h \Delta t p_{k|k-1}}{\varepsilon^2 + h^2 \Delta t p_{k|k-1}} (y_k - h(1 + b \Delta t) \hat{X}_{k-1}) \quad (6)$$

$$p_{k+1|k} = \frac{(1 + b \Delta t)^2 \varepsilon^2 p_{k|k-1}}{\varepsilon^2 + h^2 \Delta t p_{k|k-1}} + \sigma^2 \Delta t, \quad (7)$$

avec les notations:

$$p_k \triangleq E[(X_k - \hat{X}_k)^2], \quad \hat{X}_{k|k-1} \triangleq E[X_k | Y^{k-1}] \text{ et } p_{k|k-1} \triangleq E[(X_k - \hat{X}_{k|k-1})^2].$$

En outre,

$$p_k = \frac{\varepsilon^2 p_{k|k-1}}{\varepsilon^2 + h^2 \Delta t p_{k|k-1}}. \quad (8)$$

Notre objectif est de construire un processus $\{M_k\}$ qui approche $\{\hat{X}_k\}$ et qui est plus simple à calculer que ce dernier. Dans ce but on commencera par étudier le gain stationnaire, pour passer en suite à la recherche d'une approximation de ce même gain, selon les différentes valeurs de α . Le gain du filtre ayant un rapport avec la variance de l'erreur d'estimation on déduira simultanément des résultats sur cette deuxième quantité et ses approximations.

2.1 Calcul de quelques expressions utiles

Dans cette section on calcule certaines expressions qui nous seront nécessaires dans les sections suivantes.

Calcul du gain stationnaire

On commence par déterminer la covariance de l'erreur de prédiction $p_{k|k-1}$ dans une situation stationnaire, ce qui est équivalent à calculer la valeur stationnaire de la covariance de l'erreur d'estimation p_k (notée p_s^+), puisqu'on passe de l'une à l'autre par l'expression (8).

On notera θ_k le gain à l'instant t_k :

$$\theta_k \triangleq \frac{h \Delta t p_{k|k-1}}{\varepsilon^2 + h^2 \Delta t p_{k|k-1}} = \frac{h \Delta t}{\varepsilon^2} p_k.$$

Soit p_s la valeur stationnaire de $p_{k|k-1}$, i.e. la solution de

$$p_s = \frac{(1 + b \Delta t)^2 \varepsilon^2 p_s}{\varepsilon^2 + h^2 \Delta t p_s} + \sigma^2 \Delta t.$$

On obtient

$$p_s = \frac{-[\varepsilon^2 - (1 + b \Delta t)^2 \varepsilon^2 - \sigma^2 h^2 \Delta t^2] + \Delta t \rho(\varepsilon, \Delta t)}{2h^2 \Delta t},$$

où

$$\rho(\varepsilon, \Delta t) = \left[[(2b + b^2 \Delta t) \varepsilon^2 + \sigma^2 h^2 \Delta t]^2 + 4\sigma^2 h^2 \varepsilon^2 \right]^{\frac{1}{2}}.$$

Ainsi

$$p_s = \frac{(2b + b^2 \Delta t) \varepsilon^2 + \sigma^2 h^2 \Delta t + \rho(\varepsilon, \Delta t)}{2h^2}.$$

Le gain stationnaire sera donc,

$$\theta_s = \frac{h p_s \Delta t}{\varepsilon^2 + h^2 \Delta t p_s} = \frac{h \Delta t}{\varepsilon^2} p_s^+.$$

Le filtre approché

On considère un schéma qui résulte de (6) en remplaçant θ_k par sa valeur stationnaire θ_s ou, d'une façon plus générale, par une approximation de cette valeur. Dans la suite, $\bar{\theta}$ désignera donc soit θ_s soit une approximation de θ_s .

On considère le processus $\{M_k\}$ donné par l'expression

$$M_{k+1} = (1 + b \Delta t) M_k + \bar{\theta} (y_{k+1} - h(1 + b \Delta t) M_k), M_0 = m_0. \quad (9)$$

On obtient les résultats qui suivent.

Le premier résultat concerne les différences $\theta_k - \theta_s$ et $p_{k|k-1} - p_s$.

Lemme 2.2 Soit $\eta_k \triangleq \theta_k - \theta_s$ et $\mu_k \triangleq p_{k|k-1} - p_s$. On a les expressions

$$\eta_{k+1} = \frac{h \Delta t \varepsilon^2}{[\varepsilon^2 + h^2 \Delta t p_{k+1|k}][\varepsilon^2 + h^2 \Delta t p_s]} \mu_{k+1}$$

et

$$\mu_k = \left(\frac{c_0}{c_2}\right)^k \frac{c_2}{c_1 \mu_0 \sum_{i=0}^{k-1} \left(\frac{c_0}{c_2}\right)^i + c_2} \mu_0$$

où

$$\begin{aligned} c_0 &\triangleq (1 + b \Delta t)^2 \varepsilon^4 \\ c_1 &\triangleq h^2 \Delta t (\varepsilon^2 + h^2 \Delta t p_s) \\ c_2 &\triangleq (\varepsilon^2 + h^2 \Delta t p_s)^2. \end{aligned}$$

Preuve De la définition de η_k vient que

$$\begin{aligned} \eta_{k+1} &= \frac{h \Delta t p_{k+1|k}}{\varepsilon^2 + h^2 \Delta t p_{k+1|k}} - \frac{h p_s \Delta t}{\varepsilon^2 + h^2 \Delta t p_s} \\ &= \frac{h \Delta t [\varepsilon^2 p_{k+1|k} + h^2 \Delta t p_{k+1|k} p_s - \varepsilon^2 p_s - h^2 \Delta t p_{k+1|k} p_s]}{[\varepsilon^2 + h^2 \Delta t p_{k+1|k}][\varepsilon^2 + h^2 \Delta t p_s]}, \end{aligned}$$

d'où l'expression de η_{k+1} .

Quand à μ_k , on a la formule récursive

$$\begin{aligned} \mu_{k+1} &= \frac{(1 + b \Delta t)^2 \varepsilon^2 p_{k|k-1}}{\varepsilon^2 + h^2 \Delta t p_{k|k-1}} - \frac{(1 + b \Delta t)^2 \varepsilon^2 p_s}{\varepsilon^2 + h^2 \Delta t p_s} \\ &= \frac{(1 + b \Delta t)^2 \varepsilon^4}{[\varepsilon^2 + h^2 \Delta t \mu_k + h^2 \Delta t p_s][\varepsilon^2 + h^2 \Delta t p_s]} \mu_k \\ &= \frac{c_0}{c_1 \mu_k + c_2} \mu_k. \end{aligned}$$

Donc

$$\mu_{k+1} = \frac{c_0}{c_1} \frac{1}{\mu_k + \frac{c_2}{c_1}} \mu_k. \quad (10)$$

Suivant un raisonnement par récurrence on trouve l'expression:

$$\mu_k = \frac{c_0^k}{c_1 \mu_0 \sum_{i=0}^{k-1} c_2^{k-1-i} c_0^i + c_2^k} \mu_0 = \left(\frac{c_0}{c_2}\right)^k \frac{c_2}{c_1 \mu_0 \sum_{i=0}^{k-1} \left(\frac{c_0}{c_2}\right)^i + c_2} \mu_0$$

Lemme 2.3 Soit $\eta_k \triangleq \theta_k - \bar{\theta}$. On a l'égalité

$$E[(\hat{X}_{k+1} - M_{k+1})^2] = (1 + b \Delta t)^2 (1 - h \bar{\theta})^2 E[(\hat{X}_k - M_k)^2] \\ + \bar{\eta}_{k+1}^2 [h^2 p_{k+1|k} + \frac{\varepsilon^2}{\Delta t}],$$

où

$$1 - h \bar{\theta} = \frac{\varepsilon^2}{\varepsilon^2 + h^2 \Delta t p_s} + h(\theta_s - \bar{\theta}).$$

Preuve D'après les expressions (6) et (9),

$$\begin{aligned} \hat{X}_{k+1} - M_{k+1} &= (1 + b \Delta t) \hat{X}_k + \theta_{k+1}(y_{k+1} - h(1 + b \Delta t) \hat{X}_k) - (1 + b \Delta t) M_k \\ &\quad - \bar{\theta}(y_{k+1} - h(1 + b \Delta t) M_k) \\ &= (1 + b \Delta t)(\hat{X}_k - M_k) + (\theta_{k+1} - \bar{\theta}) y_{k+1} \\ &\quad - h(1 + b \Delta t)(\theta_{k+1} \hat{X}_k - \bar{\theta} M_k) \end{aligned}$$

Étant donné que $\theta_{k+1} = \bar{\theta} + (\theta_{k+1} - \bar{\theta})$, on a

$$\begin{aligned} \hat{X}_{k+1} - M_{k+1} &= (1 + b \Delta t)(1 - h \bar{\theta})(\hat{X}_k - M_k) \\ &\quad + (\theta_{k+1} - \bar{\theta})(y_{k+1} - h(1 + b \Delta t) \hat{X}_k), \end{aligned}$$

où $y_{k+1} - h(1 + b \Delta t) \hat{X}_k \triangleq \nu_k$ est l'innovation (ν_k indépendant de Y^k) et

$$E \nu_k^2 = h^2 p_{k+1|k} + \frac{\varepsilon^2}{\Delta t},$$

d'où le lemme. ■

Expressions asymptotiques de $E[(X_k - \hat{X}_k)^2]$ et de $E[(\hat{X}_k - M_k)^2]$ quand le pas de temps est ε^α

On prendra dans la suite $\Delta t = \varepsilon^\alpha$, $\alpha > 0$. Notre but est de faire une discussion sur l'ordre de grandeur de l'erreur associée au schéma (9), pour les différentes valeurs de α .

Remarque 2.4 On rappelle que :

$$p_s = \frac{(2b + b^2 \Delta t) \varepsilon^2 + \sigma^2 h^2 \Delta t + \rho_\alpha(\varepsilon, \Delta t)}{2h^2}$$

où

$$\rho_\alpha(\varepsilon) \triangleq [((2b + b^2 \Delta t) \varepsilon^2 + \sigma^2 h^2 \Delta t)^2 + 4 \sigma^2 h^2 \varepsilon^2]^{1/2}$$

Donc

$$c(\varepsilon^\alpha \vee \varepsilon) \leq p_s \leq C(\varepsilon^\alpha \vee \varepsilon)$$

Alors, puisque $\mu_0 \geq 0$ (i.e. la valeur initiale de la covariance, p_0 , est une constante indépendante de ε), des calculs précédents il vient que :

$$0 \leq \mu_k \leq \left(\frac{c_0}{c_2}\right)^k \mu_0 ,$$

où

$$\frac{c_0}{c_2} = \frac{(1 + b\varepsilon^\alpha)^2 \varepsilon^4}{[\varepsilon^2 + h^2 \varepsilon^\alpha p_s]^2} < 1,$$

i.e. $c_0 < c_2$, puisque

$$\begin{aligned} 1 - \frac{c_0}{c_2} &= \frac{[\varepsilon^2 + h^2 \varepsilon^\alpha p_s - (1 + b\varepsilon^\alpha) \varepsilon^2][\varepsilon^2 + h^2 \varepsilon^\alpha p_s + (1 + b\varepsilon^\alpha) \varepsilon^2]}{[\varepsilon^2 + h^2 \varepsilon^\alpha p_s]^2} \\ &= \frac{\varepsilon^\alpha [h^2 p_s - b\varepsilon^2][\varepsilon^2 + h^2 \varepsilon^\alpha p_s + (1 + b\varepsilon^\alpha) \varepsilon^2]}{[\varepsilon^2 + h^2 \varepsilon^\alpha p_s]^2} > 0, \end{aligned}$$

pour ε "assez petit".

On établit la proposition suivante:

Proposition 2.5 *On a l'estimation*

$$E[(X_{k+1} - \hat{X}_{k+1})^2] \leq \begin{cases} C \exp\{-c \frac{t_{k+1}}{\varepsilon}\} + C \varepsilon, & \text{si } \alpha \geq 1 \\ C \exp\{[4 \frac{t_{k+1}}{\varepsilon}(1 - \alpha) + 2(1 - \alpha)] \log \varepsilon\}, & \text{si } \alpha < 1. \end{cases}$$

Preuve

Par définition,

$$\begin{aligned} E[(X_{k+1} - \hat{X}_{k+1})^2] &\triangleq p_{k+1} = \frac{\varepsilon^2 (\mu_{k+1} + p_s)}{\varepsilon^2 + h^2 \varepsilon^2 (\mu_{k+1} + p_s)} \\ &\leq \frac{\varepsilon}{\varepsilon^2 + h^2 \varepsilon^\alpha p_s} \mu_{k+1} + \frac{\varepsilon^2 p_s}{\varepsilon^2 + h^2 \varepsilon^\alpha p_s} \\ &\leq \begin{cases} C (c_0/c_2)^{k+1} + C \varepsilon, & \text{si } \alpha \geq 1 \\ C \varepsilon^{2(1-\alpha)} (c_0/c_2)^{k+1} + C \varepsilon^{2-\alpha}, & \text{si } \alpha < 1 \end{cases} \end{aligned}$$

d'où la proposition. ■

D'autre part, du lemme 2.3, on obtient le lemme suivant:

Lemme 2.6 *On a l'estimation*

$$E[(\hat{X}_{k+1} - M_{k+1})^2] \leq \sum_{i=0}^{k+1} (1 - \bar{A})^{k+1-i} B_i ,$$

où

$$\begin{aligned} \bar{A} &\triangleq 1 - (1 + b\varepsilon^\alpha)^2 (1 - h\bar{\theta})^2 \\ B_{k+1} &\triangleq \bar{\eta}_{k+1}^2 \frac{\varepsilon^2 + h^2 \varepsilon^\alpha p_{k+1|k}}{\varepsilon^\alpha} \\ B_0 &\triangleq E[(\hat{X}_0 - M_0)^2]. \end{aligned}$$

Preuve Le lemme 2.3 devient:

$$E[(\hat{X}_{k+1} - M_{k+1})^2] = (1 - \bar{A}) E[(\hat{X}_k - M_k)^2] + B_{k+1} .$$

2.2 Filtre utilisant le gain stationnaire

On prend comme approximation du gain θ_k le gain stationnaire θ_s , i.e. $\bar{\theta} \equiv \theta_s$ (voir le schéma (9)). On déduira par la suite des résultats pour d'autres types d'approximations $\bar{\theta}$ à préciser.

Dans ce cas, on obtient la proposition suivante:

Proposition 2.7 *Pour le schéma (9) avec $\bar{\theta} = \theta_s$, on a l'estimation*

$$E[(\hat{X}_{k+1} - M_{k+1})^2] \leq \begin{cases} C \exp\{-c \frac{t_{k+1}}{\varepsilon}\} , & \text{si } \alpha \geq 1 \\ C \exp\{4 \frac{t_{k+1}}{\varepsilon} (1 - \alpha) \log \varepsilon\} , & \text{si } \alpha < 1 . \end{cases}$$

Preuve Suivant les notations du lemme 2.6, on a

$$\begin{aligned} 1 - \bar{A} &= (1 + b\varepsilon^\alpha)^2 (1 - h\theta_s)^2 \\ &= \frac{(1 + b\varepsilon^\alpha)^2 \varepsilon^4}{(\varepsilon^2 + h^2 \varepsilon^\alpha p_s)^2} \\ &\equiv \frac{c_0}{c_2} \\ B_{i+1} &= \eta_{i+1}^2 \frac{\varepsilon^2 + h^2 \varepsilon^\alpha p_{i+1|i}}{\varepsilon^\alpha} \end{aligned}$$

$$\begin{aligned}
&= \left[\frac{h \varepsilon^{\alpha+2}}{[\varepsilon^2 + h^2 \varepsilon^\alpha p_{i+1|i}][\varepsilon^2 + h^2 \varepsilon^\alpha p_s]} \mu_{i+1} \right]^2 \frac{\varepsilon^2 + h^2 \varepsilon^\alpha p_{i+1|i}}{\varepsilon^\alpha} \\
&= \frac{h^2 \varepsilon^{\alpha+4}}{[\varepsilon^2 + h^2 \varepsilon^\alpha p_{i+1|i}][\varepsilon^2 + h^2 \varepsilon^\alpha p_s]^2} \mu_{i+1}^2
\end{aligned}$$

et, puisque $p_{i+1|i} \geq p_s$, on a la majoration :

$$B_{i+1} \leq \frac{h^2 \varepsilon^{\alpha+4}}{[\varepsilon^2 + h^2 \varepsilon^\alpha p_s]^3} \mu_{i+1}^2 = B \mu_{i+1}^2 ,$$

où

$$B \triangleq \frac{h^2 \varepsilon^{\alpha+4}}{[\varepsilon^2 + h^2 \varepsilon^\alpha p_s]^3}$$

et

$$\mu_{i+1}^2 = [(1 - \rho)^2]^{i+1} H_{i+1} \mu_0^2 ,$$

où

$$H_{i+1} \triangleq \frac{c_2^2}{[c_1 \mu_0 \sum_{j=0}^i (c_0/c_2)^j + c_2]^2}$$

et H_0 est tel que

$$E[(\hat{X}_0 - M_0)^2] = B H_0 \mu_0^2 .$$

D'après ce lemme,

$$\begin{aligned}
E[(\hat{X}_{k+1} - M_{k+1})^2] &\leq (c_0/c_2) E[(\hat{X}_k - M_k)^2] + B[(c_0/c_2)^2]^{k+1} H_{k+1} \mu_0^2 \\
&\leq \sum_{i=0}^{k+1} (c_0/c_2)^{k+1-i} B[(c_0/c_2)^2]^i H_i \mu_0^2 \\
&= \mu_0^2 B (c_0/c_2)^{k+1} \sum_{i=0}^{k+1} (c_0/c_2)^i H_i . \tag{11}
\end{aligned}$$

On cherche maintenant un majorant pour la série $\left(\sum_{i=0}^k (c_0/c_2)^i H_i \right)_k$.

Or,

$$\begin{aligned}
H_{i+1} &= \frac{c_2^2}{\left[c_1 \mu_0 \frac{1 - (c_0/c_2)^{i+1}}{1 - c_0/c_2} + c_2 \right]^2} , \quad \text{avec} \quad \frac{c_0}{c_2} < 1 \\
&= \left(\frac{c_2 - c_0}{c_1 \mu_0} \right)^2 \frac{1}{\left[\left(1 + \frac{c_2 - c_0}{c_1 \mu_0} \right) - (c_0/c_2)^{i+1} \right]^2}
\end{aligned}$$

d'où

$$\sum_{i=1}^{k+1} (c_0/c_2)^i H_i = \left(\frac{c_2 - c_0}{c_1 \mu_0} \right)^2 \sum_{i=1}^{k+1} \frac{(c_0/c_2)^i}{[1 + (c_2 - c_0)/(c_1 \mu_0) - (c_0/c_2)^i]^2}$$

On commence par calculer l'ordre de grandeur de $\frac{c_2 - c_0}{c_1 \mu_0}$.

$$\frac{c_2 - c_0}{c_1 \mu_0} = \frac{(\varepsilon^2 + h^2 \varepsilon^\alpha p_s)^2 - (1 + b \varepsilon^\alpha)^2 \varepsilon^4}{\mu_0 h^2 \varepsilon^\alpha (\varepsilon^2 + h^2 \varepsilon^\alpha p_s)}$$

Soit Num le numérateur du membre à droite dans cette expression:

$$\begin{aligned} Num &\triangleq (\varepsilon^2 + h^2 \varepsilon^\alpha p_s)^2 - (1 + b \varepsilon^\alpha)^2 \varepsilon^4 \\ &= h^2 \varepsilon^\alpha (p_s - \frac{b}{h^2} \varepsilon^2) [\varepsilon^2 + h^2 \varepsilon^\alpha p_s + (1 + b \varepsilon^\alpha) \varepsilon^2] \end{aligned}$$

Or $p_s - \frac{b}{h^2} \varepsilon^2 = \mathcal{O}(\varepsilon^\alpha \vee \varepsilon)$, ce qui entraîne $Num = \mathcal{O}(\varepsilon^{4\alpha} \vee \varepsilon^{\alpha+3})$ et donc

$$\frac{c_2 - c_0}{c_1 \mu_0} = \mathcal{O}(\varepsilon^\alpha \vee \varepsilon).$$

De plus

$$\frac{c_2 - c_0}{c_1 \mu_0} > 0.$$

On peut faire la majoration suivante:

$$\begin{aligned} &\sum_{i=1}^{k+1} \frac{(c_0/c_2)^i}{[1 + (c_2 - c_0)/(c_1 \mu_0) - (c_0/c_2)^i]^2} \\ &\leq \int_0^\infty \frac{(c_0/c_2)^x}{[1 + (c_2 - c_0)/(c_1 \mu_0) - (c_0/c_2)^x]^2} dx \\ &= \frac{-1}{\log(c_0/c_2)} \frac{1}{[1 + (c_2 - c_0)/(c_1 \mu_0)](c_2 - c_0)/(c_1 \mu_0)}. \end{aligned}$$

Donc

$$\sum_{i=0}^{k+1} (c_0/c_2)^i H_i \leq \frac{c_2 - c_0}{c_1 \mu_0} \frac{-1}{\log(c_0/c_2)} \frac{1}{1 + (c_2 - c_0)/(c_1 \mu_0)} + H_0.$$

De l'expression (11), il vient alors que

$$E[(\hat{X}_{k+1} - M_{k+1})^2] \leq \mu_0^2 B(c_0/c_2)^{k+1} \left(\frac{c_2 - c_0}{c_1 \mu_0} \frac{-1}{\log(c_0/c_2)} \frac{1}{1 + (c_2 - c_0)/(c_1 \mu_0)} + H_0 \right) + H_1 \dots$$

Etant donné que

$$B = \begin{cases} \mathcal{O}(\varepsilon^{\alpha-1}), & \text{si } \alpha \geq 1 \\ \mathcal{O}(\varepsilon^{4(1-\alpha)}), & \text{si } \alpha < 1 \end{cases}$$

et

$$\frac{-1}{\log(c_0/c_2)} \leq \frac{1}{1 - (c_0/c_2)} = \begin{cases} \mathcal{O}(\varepsilon^{-(\alpha-1)}), & \text{si } \alpha \geq 1 \\ \mathcal{O}(1), & \text{si } \alpha < 1 \end{cases}$$

on obtient la majoration:

$$B \frac{-1}{\log(c_0/c_2)} \frac{1}{1 + (c_2 - c_0)/(c_1 \mu_0)} + H_0 \leq \begin{cases} C, & \text{si } \alpha \geq 1 \\ C \varepsilon^{4(1-\alpha)}, & \text{si } \alpha < 1 \end{cases} \quad (12)$$

et

$$BH_0 = \frac{1}{\mu_0^2} E[(\hat{X}_0 - M_0)^2].$$

Donc,

- si $\alpha \geq 1$, puisque $1 - c_0/c_2 = \mathcal{O}(\varepsilon^{\alpha-1})$, on a l'estimation

$$E[(\hat{X}_{k+1} - M_{k+1})^2] \leq C \exp\{-c \frac{t_{k+1}}{\varepsilon}\},$$

- si $\alpha < 1$, puisque $c_0/c_2 = \mathcal{O}(\varepsilon^{4(1-\alpha)})$, on a l'estimation

$$E[(\hat{X}_{k+1} - M_{k+1})^2] \leq C \exp\{[4(\frac{t_{k+1}}{\varepsilon} + 1)(1 - \alpha)] \log \varepsilon\}.$$

■

Remarque 2.8 On peut constater que, quel que soit $\alpha > 0$, les valeurs stationnaires de $E[(X_{k+1} - M_{k+1})^2]$ et $E[(X_{k+1} - \hat{X}_{k+1})^2]$ sont du même ordre de grandeur donc l'utilisation du filtre approché (voir le schéma (9)) est justifiée.

2.3 Filtres utilisant une approximation du gain stationnaire

On cherche maintenant les expressions plus générales qu'on obtient quand on utilise une approximation $\bar{\theta}$ du gain asymptotique θ_s .

Pour la construction de cette approximation on va d'abord considérer des approximations pour la covariance p_s . Elles conduisent naturellement à des approximations pour le gain θ_s .

On va estimer l'erreur $X_k - M_k$.

Proposition 2.9 Supposons que $p_s - \bar{p} = O(\varepsilon^m)$, $m > 1$, i.e. \bar{p} est une approximation de p_s d'ordre ε^m .

Alors,

$$E[(\hat{X}_{k+1} - M_{k+1})^2] \leq \begin{cases} C \exp\{-c \frac{t_{k+1}}{\varepsilon}\} + C \varepsilon^{2m-1}, & \text{si } \alpha \geq 1 \\ C \exp\{4 \frac{t_{k+1}}{\varepsilon} (1 - \alpha) \log \varepsilon\} + C \varepsilon^{2m+4-5\alpha}, & \text{si } \alpha < 1. \end{cases}$$

Preuve D'après le lemme 2.6, on a l'estimation

$$E[(\hat{X}_{k+1} - M_{k+1})^2] \leq \sum_{i=0}^{k+1} (1 - \bar{A})^{k+1-i} B_i,$$

où

$$1 - \bar{A} = \frac{(1 + b \varepsilon^\alpha)^2 \varepsilon^4}{[\varepsilon^2 + h^2 \varepsilon^\alpha \bar{p}]^2}$$

et

$$B_i = \bar{\eta}_i^2 \frac{\varepsilon^2 + h^2 \varepsilon^\alpha p_{i|i-1}}{\varepsilon^\alpha} \leq \bar{B} \bar{\mu}_i^2,$$

avec

$$\bar{B} \triangleq \frac{h^2 \varepsilon^{\alpha+4}}{[\varepsilon^2 + h^2 \varepsilon^\alpha p_s][\varepsilon^2 + h^2 \varepsilon^\alpha \bar{p}]^2}$$

et

$$\bar{\mu}_i = \mu_i + (p_s - \bar{p}).$$

Donc

$$\begin{aligned} E[(\hat{X}_{k+1} - M_{k+1})^2] &\leq \bar{B} \sum_{i=0}^{k+1} (1 - \bar{A})^{k+1-i} [\mu_i + C \varepsilon^m]^2 \\ &\leq 2\bar{B} \sum_{i=0}^{k+1} (1 - \bar{A})^{k+1-i} \mu_i^2 + C \varepsilon^{2m} \frac{\bar{B}}{\bar{A}}. \end{aligned}$$

En utilisant les majorations dans la démonstration de la proposition 2.7, on obtient la majoration:

$$E[(\hat{X}_{k+1} - M_{k+1})^2] \leq 2\bar{B} \mu_0^2 (1 - \bar{A})^{k+1} \sum_{i=0}^{k+1} \left[\frac{(c_0/c_2)^2}{1 - \bar{A}} \right]^i H_i + C \varepsilon^{2m} \frac{\bar{B}}{\bar{A}}.$$

Maintenant, si $p_s - \bar{p} \geq 0$ alors $1 - \bar{A} \geq c_0/c_2$; par contre, si $p_s - \bar{p} \leq 0$ alors $1 - \bar{A} \leq c_0/c_2$. Par conséquence, on a

$$E[(\hat{X}_{k+1} - M_{k+1})^2]$$

$$\begin{aligned}
&\leq 2\bar{B}\mu_0^2[(1-\bar{A}) \vee (c_0/c_2)]^{k+1} \sum_{i=0}^{k+1} (c_0/c_2)^i H_i + C \varepsilon^{2m} \frac{\bar{B}}{\bar{A}} \\
&\leq 2\bar{B}\mu_0^2[(1-\bar{A}) \vee (c_0/c_2)]^{k+1} \left(\frac{c_2 - c_0}{c_1 \mu_0} \frac{-1}{\log(c_0/c_2)} \frac{1}{1 + (c_2 - c_0)/(c_1 \mu_0)} + H_0 \right) \\
&\quad + C \varepsilon^{2m} \frac{\bar{B}}{\bar{A}}
\end{aligned}$$

Quelques calculs nous donnent:

$$\begin{aligned}
&\bar{B} \frac{(c_2 - c_0)/(c_1 \mu_0)}{1 + (c_2 - c_0)/(c_1 \mu_0)} \\
&= \frac{h^2 \varepsilon^{\alpha+4}}{[\varepsilon^2 + h^2 \varepsilon^\alpha p_s][\varepsilon^2 + h^2 \varepsilon^\alpha \bar{p}]^2} \\
&\quad \cdot \frac{(p_s - b/h^2 \varepsilon^2)[\varepsilon^2 + h^2 \varepsilon^\alpha p_s + (1 + b \varepsilon^\alpha) \varepsilon^2]}{\mu_0 (\varepsilon^2 + h^2 \varepsilon^\alpha p_s) + (p_s - b/h^2 \varepsilon^2)[\varepsilon^2 + h^2 \varepsilon^\alpha p_s + (1 + b \varepsilon^\alpha) \varepsilon^2]} \\
&= \begin{cases} \mathcal{O}(\varepsilon^{\alpha-1}), & \text{si } \alpha \geq 1 \\ \mathcal{O}(\varepsilon^{4(\alpha-1)}), & \text{si } \alpha < 1 \end{cases} \\
\bar{A} &= \frac{\varepsilon^\alpha [h^2 \bar{p} - b \varepsilon^2][\varepsilon^2 + h^2 \varepsilon^\alpha \bar{p} + (1 + b \varepsilon^\alpha) \varepsilon^2]}{[\varepsilon^2 + h^2 \varepsilon^\alpha \bar{p}]^2}, \\
&\quad \text{où } h^2 \bar{p} - b \varepsilon^2 \leq C(\varepsilon^\alpha \vee \varepsilon) \\
&= \begin{cases} \mathcal{O}(\varepsilon^{\alpha-1}), & \alpha \geq 1 \\ \mathcal{O}(1), & \alpha < 1 \end{cases} \\
\frac{\bar{B}}{\bar{A}} &= \frac{h^2 \varepsilon^4}{[\varepsilon^2 + h^2 \varepsilon^\alpha p_s][h^2 \bar{p} - b \varepsilon^2][\varepsilon^2 + h^2 \varepsilon^\alpha \bar{p} + (1 + b \varepsilon^\alpha) \varepsilon^2]} \\
&= \begin{cases} \mathcal{O}(1/\varepsilon), & \alpha \geq 1 \\ \mathcal{O}(\varepsilon^{4-5\alpha}), & \alpha < 1 \end{cases}
\end{aligned}$$

On obtient donc les estimations suivantes:

- si $\alpha \geq 1$,

$$E[(\hat{X}_{k+1} - M_{k+1})^2] \leq C \exp\left\{-c \frac{t_{k+1}}{\varepsilon}\right\} + C \varepsilon^{2m-1}$$

- si $\alpha < 1$, puisque $1 - \bar{A} \leq \mathcal{O}(\varepsilon^{4(1-\alpha)})$,

$$E[(\hat{X}_{k+1} - M_{k+1})^2] \leq C \exp\left\{4 \frac{t_{k+1}}{\varepsilon} (1 - \alpha) \log \varepsilon\right\} + C \varepsilon^{2m+4-5\alpha}$$

■

2.3.1 Approximations obtenues pour la covariance p_s

On utilise un développement limité de $\rho_\alpha(\varepsilon)$ pour construire une approximation \bar{p} de p_s et on obtient donc une approximation $\bar{\theta}$ de θ_s , par $\bar{\theta} = \frac{h \bar{p} \varepsilon^\alpha}{\varepsilon^2 + h^2 \varepsilon^\alpha \bar{p}}$.

On explicite quelques expressions possibles pour l'approximation de la variance \bar{p} :

Puisque, par hypothèse, $\sigma h > 0$, on a que

- (a) si $\alpha \geq 2$, le développement de Taylor de $\rho_\alpha(\varepsilon)$ nous donne, après quelques calculs,

$$\rho_\alpha(\varepsilon) = 2 h \sigma \varepsilon + \mathcal{O}(\varepsilon^3)$$

et donc

$$p_s = \frac{\sigma}{h} \varepsilon + \frac{b}{h^2} \varepsilon^2 + \frac{\sigma^2}{2} \varepsilon^\alpha + \mathcal{O}(\varepsilon^3).$$

Si on prend, par exemple,

1. $\boxed{\bar{p} = \sigma/h \varepsilon}$, i.e. $p_s - \bar{p} = \mathcal{O}(\varepsilon^2)$, alors

$$E[(\hat{X}_k - M_k)^2] \leq C \exp\{-c \frac{t_k}{\varepsilon}\} + C \varepsilon^3. \quad (13)$$

2. $\boxed{\bar{p} = \sigma/h \varepsilon + b/h^2 \varepsilon^2}$ si $\alpha \neq 2$, i.e. $p_s - \bar{p} = \mathcal{O}(\varepsilon^\alpha \vee \varepsilon^3)$, alors

$$E[(\hat{X}_k - M_k)^2] \leq C \exp\{-c \frac{t_k}{\varepsilon}\} + C (\varepsilon^{2\alpha-1} \vee \varepsilon^5).$$

3. $\boxed{\bar{p} = \sigma/h \varepsilon + b/h^2 \varepsilon^2 + \sigma^2/2 \varepsilon^\alpha}$ si $2 \leq \alpha < 3$, i.e. $p_s - \bar{p} = \mathcal{O}(\varepsilon^3)$, alors

$$E[(\hat{X}_k - M_k)^2] \leq C \exp\{-c \frac{t_k}{\varepsilon}\} + C \varepsilon^5.$$

- (b) si $1 < \alpha \leq 2$, du développement limité de $\rho_\alpha(\varepsilon)$ il résulte que

$$p_s = \frac{\sigma}{h} \varepsilon + \frac{\sigma^2}{2} \varepsilon^\alpha + \mathcal{O}(\varepsilon^2 \vee \varepsilon^{2\alpha-1}) \quad (14)$$

donc on peut prendre comme approximation de p_s par exemple:

1. $\boxed{\bar{p} = \sigma/h \varepsilon}$, i.e., $p_s - \bar{p} = \mathcal{O}(\varepsilon^\alpha)$ et alors

$$E[(\hat{X}_k - M_k)^2] \leq C \exp\{-c \frac{t_k}{\varepsilon}\} + C \varepsilon^{2\alpha-1}$$

2. $\boxed{\bar{p} = \sigma/h \varepsilon + \sigma^2/2 \varepsilon^\alpha}$, i.e., $p_s - \bar{p} = \mathcal{O}(\varepsilon^2 \vee \varepsilon^{2\alpha-1})$ et alors

$$E[(\hat{X}_k - M_k)^2] \leq C \exp\{-c \frac{t_k}{\varepsilon}\} + C(\varepsilon^3 \vee \varepsilon^{4\alpha-3})$$

(c) si $\alpha = 1$, le développement limité de $\rho_\alpha(\varepsilon)$ est un peu particulier:

$$\rho_\alpha(\varepsilon) = \sigma h \sqrt{4 + \sigma^2 h^2} \varepsilon + \frac{2b\sigma h}{\sqrt{4 + \sigma^2 h^2}} \varepsilon^2 + \mathcal{O}(\varepsilon^3),$$

ce qui entraîne:

$$p_s = \left(\frac{\sigma}{h} \sqrt{1 + \frac{\sigma^2 h^2}{4}} + \frac{\sigma^2}{2}\right) \varepsilon + \left(\frac{b}{h^2} + \frac{b\sigma}{2h \sqrt{1 + \sigma^2 h^2/4}}\right) \varepsilon^2 + \mathcal{O}(\varepsilon^3).$$

On propose, par exemple, les approximations suivantes:

1. $\boxed{\bar{p} = (\sigma/h \sqrt{1 + \sigma^2 h^2/4} + \sigma^2/2) \varepsilon}$, i.e. $p_s - \bar{p} = \mathcal{O}(\varepsilon^2)$ et alors

$$E[(\hat{X}_k - M_k)^2] \leq C \exp\{-c \frac{t_k}{\varepsilon}\} + C \varepsilon^3$$

2. $\boxed{\bar{p} = (\sigma/h \sqrt{1 + \sigma^2 h^2/4} + \sigma^2/2) \varepsilon + (b/h^2 + \frac{b\sigma}{2h \sqrt{1 + \sigma^2 h^2/4}}) \varepsilon^2}$,

i.e. $p_s - \bar{p} = \mathcal{O}(\varepsilon^3)$ et alors

$$E[(\hat{X}_k - M_k)^2] \leq C \exp\{-c \frac{t_k}{\varepsilon}\} + C \varepsilon^5.$$

(d) si $0 < \alpha < 1$, le développement de la fonction $\rho_\alpha(\varepsilon)$ devient

$$\rho_\alpha(\varepsilon) = \sigma^2 h^2 \varepsilon^\alpha + \mathcal{O}(\varepsilon^{2-\alpha})$$

d'où

$$p_s = \sigma^2 \varepsilon^\alpha + \mathcal{O}(\varepsilon^{2-\alpha}).$$

Donc, si on prend $\boxed{\bar{p} = \sigma^2 \varepsilon^\alpha}$, i.e. $p_s - \bar{p} = \mathcal{O}(\varepsilon^{2-\alpha})$, alors

$$E[(\hat{X}_k - M_k)^2] \leq C \exp\{4 \frac{t_{k+1}}{\varepsilon} (1 - \alpha) \log \varepsilon\} + C \varepsilon^{8-7\alpha}.$$

2.3.2 Approximations obtenues pour le gain

Puisque notre intérêt porte sur la construction de schémas du type du schéma (9), on essaie de décrire une approximation du gain stationnaire θ_s . En fait ces approximations peuvent aussi être obtenues directement d'après le développement de $\rho_\alpha(\varepsilon)$, sans passer par l'approximation de la variance.

On a

$$\theta_s = \frac{1}{h} \frac{1}{1 + D_0}$$

où

$$D_0 \triangleq \frac{2\varepsilon^{2-\alpha}}{(2b + b^2\varepsilon^\alpha)\varepsilon^2 + \sigma^2 h^2 \varepsilon^\alpha + \rho_\alpha(\varepsilon)}$$

et

$$\rho_\alpha(\varepsilon) \geq c(\varepsilon^\alpha \vee \varepsilon).$$

• *Premier cas: $\alpha > 1$*

$$D_0 \approx \frac{1}{h\sigma\varepsilon^{\alpha-1}},$$

puisque $\rho_\alpha(\varepsilon) = 2h\sigma\varepsilon + \mathcal{O}(\varepsilon^3 \vee \varepsilon^{2\alpha-1})$.

Une approximation de θ_s est donc

$$\boxed{\bar{\theta} = \sigma\varepsilon^{\alpha-1}}. \quad (15)$$

Proposition 2.10 Si $\alpha > 1$, avec le gain approché donné par (15), alors on a l'estimation

$$E[(\hat{X}_{k+1} - M_{k+1})^2] \leq C \exp\{-c \frac{t_{k+1}}{\varepsilon}\} + C(\varepsilon^{2\alpha-1} \vee \varepsilon^3)$$

Preuve

Puisque

$$\bar{\theta} = \frac{h\bar{p}\varepsilon^\alpha}{\varepsilon^2 + h^2\varepsilon^\alpha\bar{p}}$$

et, dans notre cas,

$$\bar{p} = \frac{\sigma\varepsilon}{h(1 - \sigma h\varepsilon^{\alpha-1})},$$

on obtient, en utilisant le développement limité de $\frac{1}{1 - \sigma h\varepsilon^{\alpha-1}}$,

$$\bar{p} = \frac{\sigma}{h}\varepsilon + \sigma^2\varepsilon^\alpha + \sigma^3 h\varepsilon^{2\alpha-1} + \mathcal{O}(\varepsilon^{3\alpha-2}).$$

Par conséquent

$$p_s - \bar{p} = c(\varepsilon^2 \vee \varepsilon^\alpha) \text{ et, si } \alpha > 2, p_s - \bar{p} > 0.$$

Donc, de la proposition 2.9, on déduit les estimations suivantes:

pour $1 < \alpha \leq 2$,

$$E[(\hat{X}_{k+1} - M_{k+1})^2] \leq C \exp\{-c \frac{t_{k+1}}{\varepsilon}\} + C \varepsilon^{2\alpha-1}$$

pour $\alpha \geq 2$,

$$E[(\hat{X}_{k+1} - M_{k+1})^2] \leq C \exp\{-c \frac{t_{k+1}}{\varepsilon}\} + C \varepsilon^3.$$

■

Remarque 2.11 Une fois encore on trouve des estimations de $E[(\hat{X}_{k+1} - M_{k+1})^2]$ d'ordre inférieur ou égal à celui de $E[(X_{k+1} - \hat{X}_{k+1})^2]$.

Remarque 2.12 Le schéma correspondant aux approximations (13) et (14) (i.e. qui utilise le gain $\bar{\theta} = \frac{\sigma \varepsilon^{\alpha+1}}{\varepsilon^2 + \sigma h \varepsilon^{\alpha+1}}$) n'a pas d'intérêt pratique, puisque son utilisation oblige à un calcul plus compliqué que celui du schéma qu'on vient d'obtenir alors que l'ordre de l'erreur associée reste le même.

• *Deuxième cas : $\alpha < 1$*

$$D_0 \approx \frac{2}{h^2 \sigma^2} \varepsilon^{2(1-\alpha)}.$$

Donc une approximation du gain stationnaire θ_s est

$$\boxed{\bar{\theta} = \frac{1}{h}}.$$

Le schéma (9) devient alors

$$M_{k+1} = \frac{1}{h} y_{k+1} \quad (16)$$

• *Troisième cas : $\alpha = 1$*

$$D_0 \approx \frac{2}{\sigma h \sqrt{4 + \sigma^2 h^2}} \varepsilon.$$

Quand $\varepsilon \rightarrow 0$, une approximation de θ_s est encore

$$\boxed{\bar{\theta} = \frac{1}{h}}$$

et on retrouve le schéma (16).

Proposition 2.13 Si $\alpha \leq 1$, pour un tel schéma, on a:

$$E[(\hat{X}_{k+1} - M_{k+1})^2] = \mathcal{O}(\varepsilon^{4-3\alpha})$$

et

$$E[(X_{k+1} - M_{k+1})^2] = \frac{1}{h^2} \varepsilon^{2-\alpha}.$$

Preuve

i) D'après le lemme 2.3,

$$\hat{X}_{k+1} - M_{k+1} = -\frac{\varepsilon^2}{h[\varepsilon^2 + h^2 \varepsilon^\alpha p_{k+1|k}]}(y_{k+1} - h(1 + b\varepsilon^\alpha)M_k)$$

donc

$$E[(\hat{X}_{k+1} - M_{k+1})^2] = -\frac{\varepsilon^4}{h^2 \varepsilon^\alpha [\varepsilon^2 + h^2 \varepsilon^\alpha p_{k+1|k}]}$$

d'où la première expression.

ii) D'autre part,

$$X_{k+1} - M_{k+1} = -\frac{\varepsilon}{h \varepsilon^{\alpha/2}} \bar{w}_{k+1}^2,$$

d'où la deuxième expression dans la proposition.

■

2.4 Récapitulatif

Les tableaux suivants nous donnent les vitesses de convergence à 0 de l'erreur quadratique moyenne pour les différents filtres approchés qui ont fait l'objet de cet étude, mettant en évidence la dépendance de ces filtres selon le rapport entre la variance du bruit d'observation et le pas de temps.

- Pour $\alpha > 1$

gain du filtre approché	estimation de l'erreur $E[(\hat{X}_{k+1} - M_{k+1})^2]$
θ_s (gain stationnaire)	$C e^{-ct_{k+1}/\epsilon}$
$\sigma \epsilon^{\alpha-1}$	$C(\epsilon^{2\alpha-1} \vee \epsilon^3)$
$1 < \alpha \leq 2$ $\frac{\sigma \epsilon^{\alpha+1} + \sigma^2 h/2 \epsilon^{2\alpha}}{\epsilon^2 + \sigma h \epsilon^{\alpha+1} + \sigma^2 h^2/2 \epsilon^{2\alpha}}$	$C(\epsilon^3 \vee \epsilon^{4\alpha-3})$
$\alpha > 2$ $\frac{\sigma \epsilon^{\alpha+1} + b/h \epsilon^{\alpha+2}}{\epsilon^2 + \sigma h \epsilon^{\alpha+1} + b \epsilon^{\alpha+2}}$	$C(\epsilon^{2\alpha-1} \vee \epsilon^5)$
$2 \leq \alpha < 3$ $\frac{\sigma \epsilon^{\alpha+1} + b/h \epsilon^{\alpha+2} + \sigma^2 h/2 \epsilon^{2\alpha}}{\epsilon^2 + \sigma h \epsilon^{\alpha+1} + b \epsilon^{\alpha+2} + \sigma^2 h^2/2 \epsilon^{2\alpha}}$	$C \epsilon^5$

D'autre part, on a:

$$E[(X_{k+1} - \hat{X}_{k+1})^2] \leq C e^{-ct_{k+1}/\epsilon} + C \epsilon.$$

- Pour $\alpha = 1$

gain du filtre approché	estimation de l'erreur $E[(\hat{X}_{k+1} - M_{k+1})^2]$
θ_s (gain stationnaire)	$C e^{-ct_{k+1}/\epsilon}$
$\frac{1}{h}$	$C \epsilon$
$\frac{\sigma \sqrt{1 + \sigma^2 h^2/4} + \sigma^2 h/2}{1 + h(\sigma \sqrt{1 + \sigma^2 h^2/4} + \sigma^2 h/2)}$	$C \epsilon^3$
$\frac{\sigma \sqrt{1 + \frac{\sigma^2 h^2}{4}} + \frac{\sigma^2 h}{2} + (\frac{b}{h} + \frac{b \sigma}{\sqrt{1 + \sigma^2 h^2/4}}) \epsilon}{1 + h(\sigma \sqrt{1 + \frac{\sigma^2 h^2}{4}} + \frac{\sigma^2 h}{2}) + (b + \frac{b \sigma h}{2\sqrt{1 + \sigma^2 h^2/4}}) \epsilon}$	$C \epsilon^5$

D'autre part, on a:

$$E[(X_{k+1} - \hat{X}_{k+1})^2] \leq C e^{-ct_{k+1}/\epsilon} + C \epsilon.$$

- Pour $\alpha < 1$

gain du filtre approché	estimation de l'erreur $E[(\hat{X}_{k+1} - M_{k+1})^2]$
θ_s (gain stationnaire)	$C e^{ct_{k+1}/\epsilon(1-\alpha)\log \epsilon}$
$\frac{1}{h}$	$C \epsilon^{4-3\alpha}$
$\frac{\sigma^2 h \epsilon^{2\alpha}}{\epsilon^2 + \sigma^2 h^2 \epsilon^{2\alpha}}$	$C e^{4t_{k+1}/\epsilon(1-\alpha)\log \epsilon} + C \epsilon^{8-7\alpha}$

D'autre part, on a:

$$E[(X_{k+1} - \hat{X}_{k+1})^2] \leq C e^{[4t_{k+1}/\epsilon(1-\alpha) + 2(1-\alpha)]\log \epsilon} + C \epsilon^{2-\alpha}.$$

Remarque 2.14 Des calculs analogues nous permettent aussi d'estimer directement $X_k - M_k$.

3 Cas non linéaire

On suppose maintenant que les fonctions b et h sont non linéaires.

On considère le système décrit par les équations

$$\begin{cases} X_{k+1} = X_k + b(X_k) \Delta t + \sigma \sqrt{\Delta t} w_{k+1}, & X_0 = \xi \\ y_k = h(X_k) + \frac{\varepsilon}{\sqrt{\Delta t}} \bar{w}_k \end{cases} \quad (17)$$

où

(H2') ξ est une v.a. de loi de probabilité p_0 telle que:

$$p_0 \in C^1, \int \left| \frac{p'_0}{p_0} \right|^r p_0(x) dx < \infty \quad \forall r > 0.$$

Pour des raisons liées à l'application d'une méthode de changement de probabilités, pour obtenir les estimations de l'erreur quadratique moyenne, il sera plus commode de considérer notre observation sous la forme suivante:

$$\bar{y}_{k+1} = h(X_k) + \frac{\varepsilon}{\sqrt{\Delta t}} v_{k+1}, \quad (18)$$

où $\{v_k\}$ est un bruit blanc gaussien standard indépendant de $\{w_k\}$.

Formellement, l'observation \bar{y}_{k+1} contient la même information que l'observation y_k .

On suppose que \bar{Y}^k est la tribu des observations jusqu'à l'instant t_k :

$$\bar{Y}^k = \sigma(\bar{y}_1, \bar{y}_2, \dots, \bar{y}_k),$$

i.e. $\bar{Y}^k = Y^{k-1}$.

On veut étudier la "qualité" de l'approximation

$$M_{k+1} = M_k + b(M_k) \Delta t + \bar{\theta} (\bar{y}_{k+1} - h(M_k)), \quad M_0 = m_0, \quad (19)$$

correspondante à une étape de prédiction: $\hat{X}_k = E[X_k | \bar{Y}^k]$. Le gain $\bar{\theta}$ sera donné par une expression à préciser par la suite.

Des résultats qu'on obtiendra sur l'erreur de prédiction $X_k - M_k$ on déduira immédiatement les résultats équivalents sur l'erreur de filtrage, puisque, voir le paragraphe 3.1.3, on a que

$$|\hat{X}_{k-1} - \hat{X}_k| \leq c \Delta t,$$

où $\hat{X}_{k-1} = E[X_{k-1}|\bar{Y}^k]$.

Pour la construction du filtre approché on va s'inspirer de l'étude du cas linéaire et ainsi choisir une approximation du gain. L'estimation de $X_k - M_k$ sera obtenue à partir de l'étude d'une série récurrente. En revanche, pour estimer $\hat{X}_k - M_k$, on introduira d'abord deux changements de probabilités et on obtiendra une expression asymptotique pour $X_k - M_k$ qui fait intervenir les dérivées de X , par rapport à la condition initiale. Sur cette expression, on conditionnera les différents termes par rapport à la tribu \bar{Y}^k des observations et on procédera à des estimations. Selon la valeur de α , les changements de probabilités choisis seront différents et on aboutira, évidemment, à des expressions asymptotiques différentes pour $X_k - M_k$. Dans une première section on étudie le cas $\alpha > 1$ (section 3.1) et en suite le cas $\alpha = 1$ (section 3.2). Le cas $\alpha < 1$ offrant certaines difficultés dans l'application de cette méthode de démonstration, on se limitera à faire une remarque proposant un filtre approché "performant".

Hypothèses: En plus des hypothèses (H1) et (H2') on suppose que:

(H3') h est une fonction C^3 à dérivées bornées et $\|\sigma h'\| \geq \sigma h'(x) \geq c_h > 0$, $\forall x$.

(H4) b est une fonction C^2 à dérivées bornées, $|b'(x)| \leq \|b'\|$, $\forall x$.

(H5) $\Delta t = \varepsilon^\alpha$, $\alpha \geq 1$.

Remarque 3.1 La remarque 2.1 reste valable dans le cas non linéaire, c'est à dire, pour des raisons identiques, la condition " $\sigma h'(x) > 0$, $\forall x$ " peut être remplacée sans aucune difficulté par " $\sigma h'(x) \neq 0$, $\forall x$ ". Ce qui est nécessaire est que la fonction h soit supposée injective.

3.1 Le cas $\alpha > 1$

On considère d'abord le cas où $\Delta t = \varepsilon^\alpha$, avec $\alpha > 1$.

On propose un filtre approché unidimensionnel et, pour ce filtre, on estime $X_k - M_k$ et $\hat{X}_k - M_k$.

Le filtre approché

En analogie avec le cas linéaire, on prendra $\bar{\theta} = \sigma \frac{\Delta t}{\varepsilon}$ dans le schéma (19), i.e.

$$M_{k+1} = M_k + b(M_k) \Delta t + \sigma \frac{\Delta t}{\varepsilon} (\bar{y}_{k+1} - h(M_k)), \quad M_0 = m_0. \quad (20)$$

3.1.1 Estimation de $X_k - M_k$

On obtient facilement le résultat suivant, lequel généralise le résultat dans le cas linéaire:

Proposition 3.2 *Le filtre approché M_k étant donné par (20), on a l'estimation*

$$E[(X_k - M_k)^2] \leq C \exp\{-c \frac{t_k}{\varepsilon}\} + C \varepsilon. \quad (21)$$

Preuve

On peut utiliser les développements de Taylor des fonctions b et h :

$$b(X_k) = b(M_k) + b'(\zeta_k^X)(X_k - M_k)$$

et

$$h(X_k) = h(M_k) + h'(\xi_k^X)(X_k - M_k),$$

pour obtenir l'égalité

$$\begin{aligned} X_{k+1} - M_{k+1} = & [1 - \sigma \frac{\Delta t}{\varepsilon} h'(\xi_k^X) + b'(\zeta_k^X) \Delta t] (X_k - M_k) \\ & + \sigma \sqrt{\Delta t} (w_{k+1} - v_{k+1}). \end{aligned} \quad (22)$$

D'après les hypothèses (H3), (H4) et (H5) et, vu que, d'après (H1) et (H2'), $X_k - M_k$ et $w_{k+1} - v_{k+1}$ sont indépendants, on a, pour ε "assez petit" ($\varepsilon \leq (\sigma h)^{-1/(\alpha-1)}$),

$$E[(X_{k+1} - M_{k+1})^2] \leq (1 - c_h \frac{\Delta t}{\varepsilon} + \|b'\| \Delta t)^2 E[(X_k - M_k)^2] + 2 \sigma^2 \Delta t.$$

Soit

$$A \triangleq 1 - (1 - c_h \frac{\Delta t}{\varepsilon} + \|b'\| \Delta t)^2, \quad B \triangleq 2 \sigma^2 \Delta t.$$

Avec ces notations,

$$\begin{aligned} E[(X_{k+1} - M_{k+1})^2] & \leq (1 - A)^{k+1} E[(X_0 - M_0)^2] + B \sum_{i=0}^k (1 - A)^i \\ & \leq (1 - A)^{k+1} E[(X_0 - M_0)^2] + \frac{B}{A}, \end{aligned} \quad (23)$$

puisque $1 - A < 1$.

Mais

$$c \frac{\Delta t}{\varepsilon} \leq A \leq C \left(\frac{\Delta t}{\varepsilon} \right)$$

donc

$$\frac{B}{A} = \mathcal{O}(\varepsilon)$$

et (23) devient:

$$E[(X_{k+1} - M_{k+1})^2] \leq C \exp\{-c \frac{t_{k+1}}{\varepsilon}\} + C \varepsilon$$

ou, plus précisément,

$$E[(X_k - M_k)^2] \leq C (1 - A)^k + C \varepsilon. \quad (24)$$

■

3.1.2 Estimation de $\tilde{X}_k - M_k$

On montre le théorème suivant:

Théorème 3.3 *Le schéma (20) vérifie:*

$$\tilde{X}_k - M_k = \mathcal{O}(\varepsilon^{\alpha-1/2} \vee \varepsilon)$$

au sens où

$$E[|\tilde{X}_k - M_k|] \leq C \exp\{-c \frac{t_k}{\varepsilon}\} + C (\varepsilon^{\alpha-1/2} \vee \varepsilon).$$

Preuve

Elle sera divisée en plusieurs parties, utilisant des changements de probabilités, une version discrète du Théorème de Girsanov et la dérivation par rapport à la condition initiale.

Changement de probabilités.

On introduira deux changements de probabilités (étapes (a) et (b)).

(a) Le 1^{ier} changement de probabilités affectera la loi de v .

On considère la probabilité \dot{P} définie par:

$$\left. \frac{d\dot{P}}{dP} \right|_{\mathcal{F}_k} = L_k^{-1},$$

avec

$$L_k^{-1} = \exp\left\{ \sum_{i=1}^k \left[-\frac{\sqrt{\Delta t}}{\varepsilon} h(X_{i-1}) \right] v_i - \frac{1}{2} \sum_{i=1}^k \left[-\frac{\sqrt{\Delta t}}{\varepsilon} h(X_{i-1}) \right]^2 \right\}.$$

On utilise l'équation de l'observation

$$\frac{\sqrt{\Delta t}}{\varepsilon} \bar{y}_k = v_k - \left[-\frac{\sqrt{\Delta t}}{\varepsilon} h(X_{k-1}) \right],$$

où X_{k-1} est \mathcal{F}_{k-1} -mesurable.

On obtient:

$$L_k = \exp \left\{ \frac{\Delta t}{\varepsilon^2} \left[\sum_{i=1}^k h(X_{i-1}) \bar{y}_i - \frac{1}{2} \sum_{i=1}^k h^2(X_{i-1}) \right] \right\}.$$

D'après la version discrète du Théorème de Girsanov (voir l'annexe A),

sous \dot{P} (probabilité de référence), w_k et $\frac{\sqrt{\Delta t}}{\varepsilon} \bar{y}_k$ sont des \mathcal{F}_k -bruits blancs gaussiens indépendants. (\dot{P} est équivalent à P dans chaque \mathcal{F}_k .)

(b) Le 2^{ème} changement de probabilité va affecter la loi de w .

On définit la nouvelle probabilité \tilde{P} par:

$$\frac{d\tilde{P}}{dP} \Big|_{\mathcal{F}_k} = \Lambda_k^{-1},$$

avec

$$\begin{aligned} \Lambda_k^{-1} = \exp \left\{ \sum_{i=1}^k \frac{\sqrt{\Delta t}}{\varepsilon} [h(X_{i-1}) - h(M_{i-1})] w_i \right. \\ \left. - \frac{1}{2} \sum_{i=1}^k \left[\frac{\sqrt{\Delta t}}{\varepsilon} [h(X_{i-1}) - h(M_{i-1})] \right]^2 \right\}. \end{aligned}$$

Soit $\{w_k\}$ le processus défini par

$$\tilde{w}_k = w_k - \frac{\sqrt{\Delta t}}{\varepsilon} [h(X_{k-1}) - h(M_{k-1})],$$

où $X_{k-1} - M_{k-1}$ est \mathcal{F}_{k-1} -mesurable.

Alors

$$\begin{aligned} \Lambda_k = \exp \left\{ -\frac{\sqrt{\Delta t}}{\varepsilon} \sum_{i=1}^k [h(X_{i-1}) - h(M_{i-1})] \tilde{w}_i \right. \\ \left. - \frac{\Delta t}{2\varepsilon^2} \sum_{i=1}^k [h(X_{i-1}) - h(M_{i-1})]^2 \right\} \end{aligned}$$

D'après la version discrète du Théorème de Girsanov,

Sous \tilde{P} , \tilde{w}_k et $\frac{\sqrt{\Delta t}}{\varepsilon} \tilde{y}_k$ sont des bruits blancs gaussiens indépendants et X_k vérifie l'équation:

$$X_{k+1} = \sigma \frac{\Delta t}{\varepsilon} [h(X_k) - h(M_k)] + X_k + b(X_k) \Delta t + \sigma \sqrt{\Delta t} \tilde{w}_{k+1}. \quad (25)$$

La densité de P par rapport à \tilde{P} est $L_k \Lambda_k$ donc:

$\forall v.a. \psi$ P -intégrable et \mathcal{F}_k -mesurable,

$$E[\psi | \tilde{Y}^k] = \frac{\tilde{E}[\psi L_k \Lambda_k | \tilde{Y}^k]}{\tilde{E}[L_k \Lambda_k | \tilde{Y}^k]}.$$

Soit

$$S \triangleq \log(L_k \Lambda_k).$$

Puisque

$$\begin{aligned} \tilde{w}_i &= \frac{1}{\sigma \sqrt{\Delta t}} (X_i - M_i) - \frac{1}{\sigma \sqrt{\Delta t}} (X_{i-1} - M_{i-1}) \\ &\quad - \frac{\sqrt{\Delta t}}{\sigma} [b(X_{i-1}) - b(M_{i-1})] + \frac{\sqrt{\Delta t}}{\varepsilon} [\tilde{y}_i - h(X_{i-1})] \end{aligned}$$

on a

$$\begin{aligned} S &= -\frac{1}{\sigma \varepsilon} \sum_{i=1}^k [h(X_{i-1}) - h(M_{i-1})] [(X_i - M_i) - (X_{i-1} - M_{i-1})] \\ &\quad + \frac{\Delta t}{\sigma \varepsilon} \sum_{i=1}^k [h(X_{i-1}) - h(M_{i-1})] [b(X_{i-1}) - b(M_{i-1})] \\ &\quad + \frac{\Delta t}{\varepsilon^2} \sum_{i=1}^k h(M_{i-1}) \tilde{y}_i - \frac{\Delta t}{2\varepsilon^2} \sum_{i=1}^k h^2(M_{i-1}). \end{aligned} \quad (26)$$

Remarquons que les 3^{ème} et 4^{ème} termes du second membre de cette expression sont \tilde{Y}^k -adaptés et disparaîtront donc dans la normalisation.

Dérivation par rapport à la condition initiale.

On considère les variables qui interviennent dans nos calculs comme des fonctions de X_0 (condition initiale) et des processus $\{\tilde{w}_k\}$ et $\{\tilde{y}_k\}$.

On note $\psi_k = \psi(X_0, \tilde{w}_k, \tilde{y}_k)$.

On a que X est différentiable par rapport à X_0 .

On veut dériver les expressions (25) et (26). Définissons alors les processus

$$\begin{aligned} Z_k &\triangleq \frac{\partial X_k}{\partial X_0}, \quad (Z_0 \equiv 1) \\ Z_{nk} &\triangleq \frac{Z_k}{Z_n}, \quad (Z_{0k} = Z_k) \end{aligned}$$

et fixons k .

La dérivée de X par rapport à X_0 est donnée récursivement par l'expression (27).

Lemme 3.4 Pour $n \geq k$, Z_{nk} vérifie

$$\begin{aligned} Z_{nk} &= \prod_{j=k}^{n-1} \left(1 + \sigma \frac{\Delta t}{\varepsilon} h'(X_j) + \Delta t b'(X_j) \right)^{-1} \\ Z_{kk} &= 1. \end{aligned} \quad (27)$$

Preuve

De (25) vient que:

$$X_{i+1} = \sigma \frac{\Delta t}{\varepsilon} [h(X_i) - h(M_i)] + X_i + b(X_i) \Delta t + \sigma \sqrt{\Delta t} \tilde{w}_{i+1}$$

donc, en dérivant par rapport à la condition initiale X_0 , on obtient:

$$Z_{i+1} = \left(1 + \sigma \frac{\Delta t}{\varepsilon} h'(X_i) + \Delta t b'(X_i) \right) Z_i.$$

Par récurrence, on déduit l'expression de $Z_{i,i+m}$:

$$Z_{i,i+m} = \prod_{j=i}^{i+m-1} \left(1 + \sigma \frac{\Delta t}{\varepsilon} h'(X_j) + \Delta t b'(X_j) \right).$$

Pour $n \geq k$,

$$Z_{nk} = \prod_{j=k}^{n-1} \left(1 + \sigma \frac{\Delta t}{\varepsilon} h'(X_j) + \Delta t b'(X_j) \right)^{-1}.$$

■

On a la majoration suivante:

$$Z_{nk} \leq \left(1 + c_h \frac{\Delta t}{\varepsilon} + \|b'\| \Delta t \right)^{-(n-k)} \quad (28)$$

i.e.

$$Z_{nk} \leq c \exp\left\{-c \frac{t_{n-k}}{\varepsilon}\right\}.$$

D'autre part, en dérivant (26) par rapport à X_0 , on obtient l'expression:

$$\begin{aligned} \frac{\partial}{\partial X_0} \log(L_k \Lambda_k) &= -\frac{1}{\sigma \varepsilon} \sum_{i=1}^k (X_i - M_i) h'(X_{i-1}) Z_{i-1} \\ &\quad + \frac{1}{\sigma \varepsilon} \sum_{i=1}^k (X_{i-1} - M_{i-1}) h'(X_{i-1}) Z_{i-1} \\ &\quad - \frac{1}{\sigma \varepsilon} \sum_{i=1}^k [h(X_{i-1}) - h(M_{i-1})] [Z_i - Z_{i-1} - \Delta t b'(X_{i-1}) Z_{i-1}] \\ &\quad + \frac{\Delta t}{\sigma \varepsilon} \sum_{i=1}^k [b(X_{i-1}) - b(M_{i-1})] h'(X_{i-1}) Z_{i-1}. \end{aligned}$$

Expression asymptotique pour $X_k - M_k$.

L'égalité précédente nous permettra d'écrire une expression asymptotique pour $X_k - M_k$.

On obtient, tout d'abord, l'égalité

$$\begin{aligned} &\frac{1}{\sigma \varepsilon} (X_k - M_k) h'(X_{k-1}) Z_{k-1} \\ &= \frac{1}{\sigma \varepsilon} \sum_{i=1}^k (X_{i-1} - M_{i-1}) [h'(X_{i-1}) Z_{i-1} - h'(X_{i-2}) Z_{i-2}] \\ &\quad - \frac{\Delta t}{\varepsilon^2} \sum_{i=1}^k [h(X_{i-1}) - h(M_{i-1})] h'(X_{i-1}) Z_{i-1} \\ &\quad + \frac{\Delta t}{\sigma \varepsilon} \sum_{i=1}^k [b(X_{i-1}) - b(M_{i-1})] h'(X_{i-1}) Z_{i-1} - \frac{\partial}{\partial X_0} \log(L_k \Lambda_k), \end{aligned}$$

avec $Z_{-1} \triangleq 0$.

On utilise les développements de Taylor de 2nd et 1^{er} ordre, respectivement, de h et b :

$$\begin{aligned} h(M_{i-1}) - h(X_{i-1}) &= h'(X_{i-1})(M_{i-1} - X_{i-1}) + \frac{1}{2} h''(\xi_{i-1}^X)(M_{i-1} - X_{i-1})^2 \\ b(X_{i-1}) - b(M_{i-1}) &= b'(\zeta_{i-1}^X)(X_{i-1} - M_{i-1}). \end{aligned}$$

On obtient l'égalité:

$$\begin{aligned}
& X_k - M_k \\
&= \frac{1}{h'(X_{k-1})} \sum_{i=1}^k (X_{i-1} - M_{i-1}) \left[h'(X_{i-1}) Z_{k-1,i-1} - \frac{\sigma \Delta t}{\varepsilon} h'(X_{i-1})^2 Z_{k-1,i-1} \right. \\
&\quad \left. - h'(X_{i-2}) Z_{k-1,i-2} \right] \\
&\quad + \frac{\Delta t}{h'(X_{k-1})} \sum_{i=1}^k (X_{i-1} - M_{i-1}) b'(\xi_{i-1}^X) h'(X_{i-1}) Z_{k-1,i-1} \\
&\quad + \frac{\sigma \Delta t}{2 \varepsilon h'(X_{k-1})} \sum_{i=1}^k (X_{i-1} - M_{i-1})^2 h'(X_{i-1}) h''(\xi_{i-1}^X) Z_{k-1,i-1} \\
&\quad - \frac{\sigma \varepsilon}{h'(X_{k-1})} \frac{\partial}{\partial X_0} \log(L_k \Lambda_k) Z_{k-1,0} .
\end{aligned} \tag{29}$$

Mais

• pour $i \geq 2$,

$$\begin{aligned}
& h'(X_{i-1}) Z_{k-1,i-1} - \frac{\sigma \Delta t}{\varepsilon} h'(X_{i-1})^2 Z_{k-1,i-1} - h'(X_{i-2}) Z_{k-1,i-2} \\
&= h''(X_{i-2})(X_{i-1} - X_{i-2}) Z_{k-1,i-2} + \frac{1}{2} h'''(\tilde{\xi}_{i-1}^X)(X_{i-1} - X_{i-2})^2 Z_{k-1,i-2} \\
&\quad - \frac{\sigma \Delta t}{\varepsilon} \left\{ [h''(\tilde{\xi}_{i-1}^X) + \frac{1}{2} h'(X_{i-2}) h'''(\tilde{\xi}_{i-1}^X)] (X_{i-1} - X_{i-2})^2 \right. \\
&\quad \left. + h'(X_{i-2}) h''(X_{i-2})(X_{i-1} - X_{i-2}) \right\} Z_{k-1,i-2} \\
&\quad + \left[\Delta t h'(X_{i-1}) b'(X_{i-2}) \left(1 - \frac{\sigma \Delta t}{\varepsilon} h'(X_{i-1}) \right) - \frac{\sigma^2 \Delta t^2}{\varepsilon^2} h'(X_{i-1})^2 h'(X_{i-2}) \right] Z_{k-1,i-2} ,
\end{aligned}$$

puisque

$$h'(X_{i-1}) - h'(X_{i-2}) = h''(X_{i-2})(X_{i-1} - X_{i-2}) + \frac{1}{2} h'''(\tilde{\xi}_{i-1}^X)(X_{i-1} - X_{i-2})^2$$

et

$$\begin{aligned}
& h'(X_{i-1}) [h'(X_{i-1}) - h'(X_{i-2})] \\
&= h''(\tilde{\xi}_{i-1}^X)(X_{i-1} - X_{i-2})^2 + h'(X_{i-2}) [h''(X_{i-2})(X_{i-1} - X_{i-2}) \\
&\quad + \frac{1}{2} h'''(\tilde{\xi}_{i-2}^X)(X_{i-1} - X_{i-2})^2] .
\end{aligned}$$

Soit

$$\phi_i \triangleq \Delta t h'(X_{i-1}) b'(X_{i-2}) \left(1 - \frac{\sigma \Delta t}{\varepsilon} h'(X_{i-1}) \right) - \frac{\sigma^2 \Delta t^2}{\varepsilon^2} h'(X_{i-1})^2 h'(X_{i-2}) .$$

On a la majoration

$$|\phi_i| \leq C \left(\frac{\Delta t^2}{\varepsilon^2} \vee \Delta t \right), \forall i \geq 2.$$

• Pour $i = 1$,

$$\begin{aligned} & h'(X_{i-1}) Z_{k-1,i-1} - \frac{\sigma \Delta t}{\varepsilon} h'^2(X_{i-1}) Z_{k-1,i-1} - h'(X_{i-2}) Z_{k-1,i-2} \\ &= (1 - \sigma \frac{\Delta t}{\varepsilon} h'(X_0)) h'(X_0) Z_{k-1,0} \\ &= \phi_1 Z_{k-1,0} \end{aligned}$$

où

$$\phi_1 \triangleq h'(X_0) (1 - \sigma \frac{\Delta t}{\varepsilon} h'(X_0)).$$

L'égalité (29) devient:

$$\begin{aligned} & X_k - M_k \\ &= \frac{1}{h'(X_{k-1})} \sum_{i=2}^k (X_{i-1} - M_{i-1}) (X_{i-1} - X_{i-2}) h''(X_{i-2}) Z_{k-1,i-2} \\ &+ \frac{1}{2 h'(X_{k-1})} \sum_{i=2}^k (X_{i-1} - M_{i-1}) (X_{i-1} - X_{i-2})^2 h'''(\xi_{i-1}^X) Z_{k-1,i-2} \\ &- \frac{\sigma \Delta t}{\varepsilon h'(X_{k-1})} \sum_{i=2}^k (X_{i-1} - M_{i-1}) \left\{ [h''^2(\xi_{i-1}^X) + \frac{1}{2} h'(X_{i-2}) h'''(\xi_{i-1}^X)] (X_{i-1} - X_{i-2})^2 \right. \\ &\quad \left. + h'(X_{i-2}) h''(X_{i-2}) (X_{i-1} - X_{i-2}) \right\} Z_{k-1,i-2} \\ &+ \frac{1}{h'(X_{k-1})} \sum_{i=2}^k (X_{i-1} - M_{i-1}) \left[\frac{\phi_i}{1 + \sigma \Delta t / \varepsilon h'(X_{i-2}) + \Delta t b'(X_{i-2})} \right. \\ &\quad \left. + \Delta t b'(\zeta_{i-1}^X) h'(X_{i-1}) \right] Z_{k-1,i-1} \\ &+ \frac{1}{h'(X_{k-1})} (X_0 - M_0) h'(X_0) (1 - \frac{\sigma \Delta t}{\varepsilon} h'(X_0) + \Delta t b'(\zeta_0^X)) Z_{k-1,0} \\ &+ \frac{\sigma \Delta t}{2 \varepsilon h'(X_{k-1})} (X_0 - M_0)^2 h'(X_0) h''(\zeta_0^X) Z_{k-1,0} \\ &+ \frac{\sigma \Delta t}{2 \varepsilon h'(X_{k-1})} \sum_{i=2}^k (X_{i-1} - M_{i-1})^2 h'(X_{i-1}) h''(\xi_{i-1}^X) Z_{k-1,i-1} \\ &- \frac{\sigma \varepsilon}{h'(X_{k-1})} \frac{\partial}{\partial X_0} \log(L_k \Lambda_k) Z_{k-1,0}. \end{aligned} \tag{30}$$

On remarque que:

$$\left| \frac{\phi_i}{1 + \sigma \Delta t / \varepsilon h'(X_{i-2}) + \Delta t b'(X_{i-2})} + \Delta t b'(\zeta_{i-1}^X) h'(X_{i-1}) \right| \leq C \left(\frac{\Delta t^2}{\varepsilon^2} \vee \Delta t \right) \triangleq \rho, \forall i \geq 2.$$

On cherche maintenant des majorations dans L^1 pour les différents termes de droite dans (30), à l'exception du dernier.

Voyons d'abord que:

$$E[|(X_{i-1} - M_{i-1})(X_{i-1} - X_{i-2})|] \leq C \sqrt{\Delta t} (E[(X_{i-1} - M_{i-1})^2])^{1/2}.$$

Puisque, voir (24),

$$E[|X_{i-1} - M_{i-1}|] \leq C [(1 - A)^{1/2}]^{i-1} + C \sqrt{\varepsilon},$$

où

$$1 - A \triangleq (1 - c_h \frac{\Delta t}{\varepsilon} + \|b'\| \Delta t)^2$$

et, en outre, d'après (28),

$$Z_{nk} \leq (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-(n-k)},$$

on a que:

- le 1^{ier} terme est majoré (dans L^1) par:

$$C \sqrt{\varepsilon} \exp\{-c \frac{t_k}{\varepsilon}\} + C \varepsilon, \text{ pour } \Delta t = \varepsilon^\alpha.$$

Cette estimation est justifiée dans l'annexe B.

- le 2^{me} terme est majoré (dans L^1) par:

$$\begin{aligned} & C \Delta t \sum_{i=2}^k [(1 - c_h \frac{\Delta t}{\varepsilon} + \|b'\| \Delta t)^{i-1} + C \sqrt{\varepsilon}] (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-(k-i+1)} \\ & \leq C \Delta t (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-k} \sum_{i=1}^{k-1} [1 - (c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^2]^i \\ & \quad + C \Delta t \sqrt{\varepsilon} \sum_{i=1}^{k-1} (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-i} \\ & \leq C \Delta t (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-k} (k-1) \\ & \quad + C \Delta t \sqrt{\varepsilon} \frac{(1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-1}}{1 - (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-1}} \\ & \leq C \varepsilon \exp\{-c \frac{t_k}{\varepsilon}\} + C \varepsilon^{3/2}, \end{aligned}$$

puisque

$$E[|X_{i-1} - X_{i-2}|(X_{i-1} - M_{i-1})^2] \leq C \Delta t (E[(X_{i-1} - M_{i-1})^2])^{1/2}.$$

- le 3^{me} terme est majoré (dans L^1) par:

$$C \varepsilon^{\alpha-1/2} \exp\{-c \frac{t_k}{\varepsilon}\} + C \varepsilon^\alpha ,$$

puisque'il s'agit de la somme d'un terme du même type du 1^{ier} terme ici étudié avec un terme du même type du 2^{me}, multipliés par un facteur d'ordre $\frac{\Delta t}{\varepsilon}$.

- le 4^{me} terme est majoré (dans L^1) par:

$$\begin{aligned} & C \rho \sum_{i=2}^k [(1 - c_h \frac{\Delta t}{\varepsilon} + \|b'\| \Delta t)^{i-1} + C \sqrt{\varepsilon}] (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-(k-i)} \\ & \leq C (\varepsilon \vee \varepsilon^{\alpha-1}) \exp\{-c \frac{t_{k-1}}{\varepsilon}\} + C (\varepsilon^{\alpha-1/2} \vee \varepsilon^{3/2}) \end{aligned}$$

- le 5^{me} et 6^{me} termes sont majorés (dans L^1) par:

$$C (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-(k-1)} \leq C \exp\{-c \frac{t_{k-1}}{\varepsilon}\}$$

- le 7^{me} terme est majoré (dans L^1) par:

$$\begin{aligned} & C \frac{\Delta t}{\varepsilon} \sum_{i=2}^k \{(1 - c_h \frac{\Delta t}{\varepsilon} + \|b'\| \Delta t)^{2(i-1)} + C \varepsilon\} (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-(k-i)} \\ & \leq C \exp\{-c \frac{t_k}{\varepsilon}\} + C \varepsilon \end{aligned}$$

donc on établit l'estimation

$$X_k - M_k + \frac{\sigma \varepsilon}{h'(X_{k-1})} \frac{\partial}{\partial X_0} \log(L_k \Lambda_k) Z_{k-1,0} = \mathcal{O}(\varepsilon^{\alpha-1/2} \vee \varepsilon) , \quad (31)$$

au sens où

$$E[|X_k - M_k + \frac{\sigma \varepsilon}{h'(X_{k-1})} \frac{\partial}{\partial X_0} \log(L_k \Lambda_k) Z_{k-1,0}|] \leq C \exp\{-c \frac{t_{k-1}}{\varepsilon}\} + C (\varepsilon^{\alpha-1/2} \vee \varepsilon) .$$

Vu que h' est bornée supérieurement et inférieurement, on peut écrire que

$$\frac{h'(X_{k-1})}{h'(M_{k-1})} (X_k - M_k) + \frac{\sigma \varepsilon}{h'(M_{k-1})} \frac{\partial}{\partial X_0} \log(L_k \Lambda_k) Z_{k-1,0} = \mathcal{O}(\varepsilon^{\alpha-1/2} \vee \varepsilon) .$$

On utilise la décomposition

$$X_k - M_k = \frac{h'(X_{k-1})}{h'(M_{k-1})} (X_k - M_k) + \left(1 - \frac{h'(X_{k-1})}{h'(M_{k-1})}\right) (X_k - M_k),$$

où le 2nd terme est un produit de deux facteurs d'ordre $\mathcal{O}(\sqrt{\varepsilon})$ dans L^2 (cf. proposition 3.2) et, par conséquent, d'ordre $\mathcal{O}(\varepsilon)$ dans L^1 .

Prenons l'espérance conditionnelle par rapport à \bar{Y}^k .

On va montrer que:

$$E\left[\frac{\partial}{\partial X_0} \log(L_k \Lambda_k) Z_{k-1,0} | \bar{Y}^k\right] = \mathcal{O}(1) \quad (32)$$

et alors on a

$$\hat{X}_k - M_k = \mathcal{O}(\varepsilon^{\alpha-1/2} \vee \varepsilon),$$

au sens où

$$E[|\hat{X}_k - M_k|] \leq C \exp\left\{-c \frac{t_k}{\varepsilon}\right\} + C(\varepsilon^{\alpha-1/2} \vee \varepsilon).$$

On utilise le lemme 3.5 pour prouver ce résultat.

Lemme 3.5 *On a l'estimation*

$$E\left[\frac{\partial}{\partial X_0} \log(L_k \Lambda_k) Z_{k-1,0} | \bar{Y}^k\right] = \mathcal{O}(1),$$

au sens où

$$\left| E\left[\frac{\partial}{\partial X_0} \log(L_k \Lambda_k) Z_{k-1,0} | \bar{Y}^k\right] \right| \leq C.$$

Preuve

On rappelle d'abord que:

Si $g \in C^1$ et g et $\frac{\partial g}{\partial x}$ sont des fonctions intégrables par rapport à la mesure de Lebesgue, alors

$$\int \frac{\partial g}{\partial x} dx = 0. \quad (33)$$

En particulier,

soit ψ une v.a. différentiable par rapport à X_0 telle que:

$$\tilde{E}[|\psi| + |\psi \frac{p'_0}{p_0}(X_0)| + |\frac{\partial \psi}{\partial X_0}|] < \infty, \quad (34)$$

soit

$$g(x) = p_0(x) \psi(x, \{\tilde{w}_k\}, \{\tilde{y}_k\}),$$

où $\psi(x, \{\tilde{w}_k\}, \{\tilde{y}_k\})$ est une fonction différentiable par rapport à x ,
alors

$$\tilde{E}[\frac{\partial \psi}{\partial X_0} + \psi \frac{p'_0}{p_0}(X_0) | \tilde{w}, \tilde{y}] = 0. \quad (35)$$

On applique la formule de changement de probabilité,

$$\begin{aligned} E[\frac{\partial}{\partial X_0} \log(L_k \Lambda_k) Z_{k-1,0} | \bar{Y}^k] &= \frac{\tilde{E}[\frac{\partial}{\partial X_0} \log(L_k \Lambda_k) Z_{k-1,0} | \bar{Y}^k]}{\tilde{E}[L_k \Lambda_k | \bar{Y}^k]} \\ &= \frac{\tilde{E}[\frac{\partial}{\partial X_0} (L_k \Lambda_k) Z_{k-1,0} | \bar{Y}^k]}{\tilde{E}[L_k \Lambda_k | \bar{Y}^k]}. \end{aligned}$$

Soit

$$\tilde{\phi}_k \triangleq \tilde{E}[\frac{\partial}{\partial X_0} (L_k \Lambda_k) Z_{k-1,0} | \bar{Y}^k].$$

Puisque

$$\frac{\partial}{\partial X_0} (L_k \Lambda_k Z_{k-1,0}) = \frac{\partial}{\partial X_0} (L_k \Lambda_k) Z_{k-1,0} + (L_k \Lambda_k) \frac{\partial}{\partial X_0} Z_{k-1,0}$$

on a que

$$\tilde{\phi}_k = \tilde{E}[\frac{\partial}{\partial X_0} (L_k \Lambda_k Z_{k-1,0}) | \bar{Y}^k] - \tilde{E}[(L_k \Lambda_k) \frac{\partial}{\partial X_0} Z_{k-1,0} | \bar{Y}^k].$$

Étant donné que $(L_k \Lambda_k) Z_{k-1,0}$ vérifie la condition d'intégrabilité (34) on applique (35) pour obtenir:

$$\tilde{E}[\frac{\partial}{\partial X_0} (L_k \Lambda_k Z_{k-1,0}) | \bar{Y}^k] = -\tilde{E}[(L_k \Lambda_k) Z_{k-1,0} \frac{p'_0}{p_0}(X_0) | \bar{Y}^k]$$

et alors

$$\tilde{\phi}_k = -\tilde{E}[\frac{p'_0}{p_0}(X_0) (L_k \Lambda_k) Z_{k-1,0} | \bar{Y}^k] - \tilde{E}[(L_k \Lambda_k) \frac{\partial}{\partial X_0} Z_{k-1,0} | \bar{Y}^k].$$

On a donc

$$\begin{aligned}
& E\left[\frac{\partial}{\partial X_0} \log(L_k \Lambda_k) Z_{k-1,0} | \bar{Y}^k\right] \\
&= -\frac{\tilde{E}[(L_k \Lambda_k) \frac{p'_0}{p_0}(X_0) Z_{k-1,0} | \bar{Y}^k]}{\tilde{E}[L_k \Lambda_k | \bar{Y}^k]} - \frac{\tilde{E}[(L_k \Lambda_k) \frac{\partial}{\partial X_0} Z_{k-1,0} | \bar{Y}^k]}{\tilde{E}[L_k \Lambda_k | \bar{Y}^k]} \\
&= -E\left[\frac{p'_0}{p_0}(X_0) Z_{k-1,0} | \bar{Y}^k\right] - E\left[\frac{\partial}{\partial X_0} Z_{k-1,0} | \bar{Y}^k\right].
\end{aligned}$$

Puisqu'on avait supposé que

$$\int \left| \frac{p'_0}{p_0}(x) \right|^r p_0(x) dx < \infty, \quad p_0 \in C^1 \quad (36)$$

et, de (28), on a la majoration

$$\begin{aligned}
Z_{k-1,0} &\leq (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-(k-1)} \\
&\leq C \exp\left\{-c \frac{t_{k-1}}{\varepsilon}\right\}
\end{aligned}$$

il vient que le 1^{ier} terme est d'ordre $C \exp\left\{-c \frac{t_{k-1}}{\varepsilon}\right\}$. Alors, il nous suffit de démontrer que

$$E\left[\frac{\partial}{\partial X_0} Z_{k-1,0} | \bar{Y}^k\right] = \mathcal{O}(1). \quad (37)$$

Or, de (27), on peut obtenir la dérivée partielle de Z_{n0} par rapport à X_0 :

$$\begin{aligned}
\frac{\partial}{\partial X_0} Z_{n0} &= -\sum_{i=0}^{n-1} \left(\sigma \frac{\Delta t}{\varepsilon} h''(X_i) + \Delta t b''(X_i) \right) \left(1 + \sigma \frac{\Delta t}{\varepsilon} h'(X_i) + \Delta t b'(X_i) \right)^{-1} \\
&\quad \cdot \prod_{j=i}^{n-1} \left(1 + \sigma \frac{\Delta t}{\varepsilon} h'(X_j) + \Delta t b'(X_j) \right)^{-1}
\end{aligned}$$

et

$$\frac{\partial}{\partial X_0} Z_{0,0} = 0$$

donc

$$\left| \frac{\partial}{\partial X_0} Z_{n0} \right| \leq \left(\frac{\Delta t}{\varepsilon} \|\sigma h''\| + \|b''\| \Delta t \right) (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-1} \sum_{i=1}^n (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-i}$$

$$\begin{aligned}
&\leq C \frac{\Delta t}{\varepsilon} (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-1} \frac{(1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-1}}{1 - (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-1}} \\
&\leq C,
\end{aligned}$$

ce qui prouve (37) et termine la démonstration du lemme. ■

Remarque 3.6 On peut démontrer un lemme plus général que celui-là, lequel peut être énoncé de la façon suivante:

Lemme 3.7 Soit $\{F\}_k$ un processus adapté (dépendant de ε), différentiable par rapport à X_0 ($\forall k = 0, 1, \dots, K$).

Si les moments de $\{\frac{\partial F}{\partial X_0}\}_k$ sont finis et si $F_k = \mathcal{O}(\varepsilon^q)$, alors, pour un certain $q \geq 0$,

$$E[F_k \frac{\partial}{\partial X_0} \log(L_k \Lambda_k) Z_{k-1,0} | \bar{Y}^k] = -E[\frac{\partial F_k}{\partial X_0} Z_{k-1,0} | \bar{Y}^k] + \mathcal{O}(\varepsilon^q). \quad (38)$$

3.1.3 Estimation de l'erreur de filtrage

Enfin, pour obtenir une estimation de l'erreur commise dans une étape de filtrage, il nous suffit de remarquer qu'une approximation du filtre optimal $\hat{X}_{k-1} = E[X_{k-1} | \bar{Y}^k]$ est donnée par le schéma (20), étant donné que

$$\begin{aligned}
\tilde{X}_k - \hat{X}_{k-1} &= E[X_k | \bar{Y}^k] - E[X_{k-1} | \bar{Y}^k] \\
&= E[(b(X_{k-1}) \Delta t + \sigma \sqrt{\Delta t} w_k) | \bar{Y}^k] \\
&= \Delta t E[b(X_{k-1}) | \bar{Y}^k],
\end{aligned}$$

i.e.

$$|\hat{X}_{k-1} - \tilde{X}_k| \leq c \Delta t. \quad (39)$$

Corollaire 3.8 Le schéma (20) vérifie:

$$E[|\hat{X}_k - M_k|] \leq C \exp\{-c \frac{t_k}{\varepsilon}\} + C(\varepsilon^{\alpha-1/2} \vee \varepsilon).$$

Remarque 3.9 Dans le cas h linéaire, on voit immédiatement qu'il ne reste plus que les 4^{me}, 5^{me} et 8^{me} termes dans l'expression (30), tous les autres disparaissent et, dans le lemme 3.5 au lieu de $\mathcal{O}(1)$ on a $\mathcal{O}(\varepsilon)$. On obtient ainsi que, si h est linéaire,

$$\hat{X}_k - M_k = \mathcal{O}(\varepsilon^{\alpha-1/2} \vee \varepsilon^{3/2}),$$

i.e.

$$E[|\hat{X}_k - M_k|] \leq C \exp\{-c \frac{t_k}{\varepsilon}\} + C(\varepsilon^{\alpha-1/2} \vee \varepsilon^{3/2}).$$

3.2 Le cas $\alpha = 1$

On considère maintenant le cas où le paramètre Δt vaut ε , i.e. $\alpha = 1$ dans l'hypothèse (H5).

Le système considéré est donc décrit par les équations:

$$\begin{cases} X_{k+1} = X_k + b(X_k)\varepsilon + \sigma\sqrt{\varepsilon}w_{k+1}, & X_0 = \xi \\ \bar{y}_{k+1} = h(X_k) + \sqrt{\varepsilon}v_{k+1}. \end{cases} \quad (40)$$

En plus des hypothèses énoncées en début de cette section dédiée au cas non linéaire, on suppose que:

$$(H6) \quad \|\sigma h'\| < 2c_h.$$

Remarque 3.10 Une telle restriction signifie que notre étude n'est valable que si la fonction h "n'est pas trop éloignée" d'une fonction linéaire.

Remarque 3.11 En fait on peut considérer une hypothèse légèrement moins restrictive, mais moins intuitive quand à l'interprétation. Il suffit, dans le lemme 3.13, d'avoir le soin de chercher une minoration du gain un peu plus fine.

Le filtre approché

On aimerait pouvoir construire un filtre approché unidimensionnel, pour ce problème. Cependant cette tâche se révèle délicate, ce qui nous ramène à la recherche d'un filtre du genre du filtre de Kalman étendu.

On considère le filtre bidimensionnel suivant:

$$M_{k+1} = M_k + b(M_k)\varepsilon + \bar{\theta}_k(\bar{y}_{k+1} - h(M_k)), \quad M_0 = m_0, \quad (41)$$

et

$$\bar{\theta}_k = \frac{h'(M_k)[\bar{\theta}_{k-1} + \sigma^2 h'(M_{k-1})]}{h'^2(M_k)[\bar{\theta}_{k-1} + \sigma^2 h'(M_{k-1})] + h'(M_{k-1})}, \quad \bar{\theta}_0 = \theta_0, \quad (42)$$

avec m_0 quelconque et θ_0 tel que $\sigma \theta_0 > 0$.

Pour b et h linéaires, ce filtre correspond à prendre le filtre de Kalman - Bucy et négliger les termes en $b \Delta t$ dans l'expression du gain.

On cherche d'abord à estimer $X_k - M_k$ et ensuite $\hat{X}_k - M_k$. Les méthodes utilisées pour obtenir ces estimations sont essentiellement les mêmes que dans la section précédente. La même considération sur l'erreur de filtrage qui a été faite dans le cas précédent est vraie, i.e., avec les notations du paragraphe 3.1.3, on s'intéresse d'abord à l'estimation de $\hat{X}_k - M_k$ et puis on utilise le fait que

$$|\hat{X}_{k-1} - \tilde{X}_k| \leq c \varepsilon$$

pour déduire le résultat recherché, l'estimation de l'erreur de filtrage.

3.2.1 Estimation de $X_k - M_k$

On établit la proposition suivante:

Proposition 3.12 *On a l'estimation*

$$E[(X_k - M_k)^2] \leq C \exp\left\{-c \frac{t_k}{\varepsilon}\right\} + C \varepsilon.$$

Pour démontrer cette proposition on fera appel à un lemme (lemme 3.13) qu'on présentera par la suite.

Preuve

Etant données les hypothèses de régularité sur les fonctions h et b , on peut écrire leurs développements de Taylor du 1^{ier} ordre:

$$h(X_k) - h(M_k) = h'(\xi_k^X)(X_k - M_k)$$

et

$$b(X_k) - b(M_k) = b'(\zeta_k^X)(X_k - M_k).$$

On obtient alors l'expression récursive:

$$\begin{aligned} X_{k+1} - M_{k+1} &= (X_k - M_k) + [b(X_k) - b(M_k)] \varepsilon - \bar{\theta}_k [h(X_k) - h(M_k)] \\ &\quad + \sqrt{\varepsilon} (\sigma w_{k+1} - \bar{\theta}_k v_{k+1}) \\ &= [1 - \bar{\theta}_k h'(\xi_k^X) + b'(\zeta_k^X) \varepsilon] (X_k - M_k) \\ &\quad + \sqrt{\varepsilon} (\sigma w_{k+1} - \bar{\theta}_k v_{k+1}). \end{aligned}$$

D'après le lemme 3.13 ci-dessous, pour ε_0 "assez petit", on a

$$\forall \varepsilon \leq \varepsilon_0, |1 - \bar{\theta}_k h'(\xi_k^X) + b'(\zeta_k^X) \varepsilon| \leq 1 - A < 1.$$

D'autre part, puisque $(\bar{\theta}_k)_k$ est une suite bornée, on a la majoration

$$E[(\sigma w_{k+1} - \bar{\theta}_k v_{k+1})^2] = \sigma^2 + E[\bar{\theta}_k^2] + 2\sigma E[\bar{\theta}_k w_{k+1} v_{k+1}] \leq C$$

et

$$E\left\{(\sigma w_{k+1} - \bar{\theta}_k v_{k+1}) \left[(X_k - M_k) + [b(X_k) - b(M_k)] \varepsilon - \bar{\theta}_k [h(X_k) - h(M_k)]\right]\right\} = 0$$

donc

$$\begin{aligned} E[(X_{k+1} - M_{k+1})^2] &\leq (1 - A)^2 E[(X_k - M_k)^2] + C \varepsilon \\ &\leq (1 - A)^{2(k+1)} E[(X_0 - M_0)^2] + C \frac{\varepsilon}{A}, \end{aligned}$$

d'où la proposition. ■

Il nous reste à établir le lemme

Lemme 3.13 *Il existe $\varepsilon_0 > 0$ tel que, pour un certain $A > 0$,*

$$\forall \varepsilon \leq \varepsilon_0, |1 - \bar{\theta}_k h'(\xi_k^X) + b'(\zeta_k^X) \varepsilon| \leq 1 - A.$$

Preuve

On utilise le fait que $\bar{\theta}_{k-1}/\sigma > 0$ et l'hypothèse (H3'). De (42) il vient que

$$\begin{aligned} \frac{\bar{\theta}_k}{\sigma} &= \frac{\sigma h'(M_k) [\bar{\theta}_{k-1}/\sigma + \sigma^2 h'(M_{k-1})]}{\sigma^2 h'^2(M_k) [\bar{\theta}_{k-1}/\sigma + \sigma^2 h'(M_{k-1})] + \sigma h'(M_{k-1})} \\ &< \frac{1}{\sigma h'(M_k)} \leq \frac{1}{c_h} \end{aligned}$$

et alors on a que

$$\begin{aligned} \frac{\bar{\theta}_k}{\sigma} &> \frac{\sigma^2 h'(M_k) h'(M_{k-1})}{\sigma^2 h'^2(M_k) [1/(\sigma h'(M_{k-1})) + \sigma h'(M_{k-1})] + \sigma h'(M_{k-1})} \\ &\geq \frac{c_h^3}{2 \|\sigma h'\|^2 + \|\sigma h'\|^4}. \end{aligned}$$

On obtient les bornes inférieure et supérieure suivantes, pour $\bar{\theta}_k h'(\xi_k^X)$:

$$\frac{c_h^4}{2 \|\sigma h'\|^2 + \|\sigma h'\|^4} < \bar{\theta}_k h'(\xi_k^X) < \frac{\|\sigma h'\|}{c_h},$$

d'où on déduit que:

- pour $A_1 < \frac{c_h^4}{2\|\sigma h'\|^2 + \|\sigma h'\|^4}$ et $\varepsilon < \varepsilon_1 = \frac{1}{\|b'\|} \left(\frac{c_h^4}{2\|\sigma h'\|^2 + \|\sigma h'\|^4} - A_1 \right)$, on a la majoration

$$1 - \bar{\theta}_k h'(\xi_k^X) + b'(\zeta_k^X) \varepsilon < 1 - \frac{c_h^4}{2\|\sigma h'\|^2 + \|\sigma h'\|^4} + \|b'\| \varepsilon \leq 1 - A_1,$$

- pour $A_2 < \frac{\bar{c}}{c_h}$ et $\varepsilon < \varepsilon_2 = \frac{\bar{c} - A_2 c_h}{c_h \|b'\|}$, on a la minoration

$$1 - \bar{\theta}_k h'(\xi_k^X) + b'(\zeta_k^X) \varepsilon > 1 - \frac{\|\sigma h'\|}{c_h} - \|b'\| \varepsilon \geq A_2 - 1 > -1.$$

Il suffit de considérer

$$A = \min(A_1, A_2) \text{ et } \varepsilon_0 = \min(\varepsilon_1, \varepsilon_2).$$

■

3.2.2 Estimation de $\tilde{X}_k - M_k$

Le théorème suivant nous donne une estimation sur l'erreur de prédiction $\tilde{X}_k - M_k$.

Théorème 3.14 *On a l'estimation*

$$E[|\tilde{X}_k - M_k|] \leq C \exp\left\{-c \frac{t_k}{\varepsilon}\right\} + C \varepsilon.$$

Preuve

La méthode suivie pour démontrer ce théorème est essentiellement la même que dans le cas précédent.

Changement de probabilités

- On considère la probabilité \tilde{P} précédemment définie, avec $\Delta t = \varepsilon$.
- On considère la nouvelle probabilité \tilde{P} définie par:

$$\frac{d\tilde{P}}{dP}\bigg|_{\mathcal{F}_k} = \Lambda_k^{-1},$$

avec

$$\Lambda_k^{-1} = \exp \left\{ \frac{\sigma}{\sqrt{\varepsilon}} \sum_{i=1}^k \frac{1}{\bar{\theta}_{i-1}} [h(X_{i-1}) - h(M_{i-1})] w_i - \frac{\sigma^2}{2\varepsilon} \sum_{i=1}^k \frac{1}{\bar{\theta}_{i-1}^2} [h(X_{i-1}) - h(M_{i-1})]^2 \right\}.$$

Soit

$$\tilde{w}_k = w_k - \frac{\sigma}{\sqrt{\varepsilon} \bar{\theta}_{k-1}} [h(X_{k-1}) - h(M_{k-1})].$$

Alors

$$\Lambda_k = \exp \left\{ -\frac{\sigma}{\sqrt{\varepsilon}} \sum_{i=1}^k \frac{1}{\bar{\theta}_{i-1}} [h(X_{i-1}) - h(M_{i-1})] \tilde{w}_i - \frac{\sigma^2}{2\varepsilon} \sum_{i=1}^k \frac{1}{\bar{\theta}_{i-1}^2} [h(X_{i-1}) - h(M_{i-1})]^2 \right\}.$$

Sous \tilde{P} , \tilde{w}_k et $\frac{1}{\sqrt{\varepsilon}} \tilde{y}_k$ sont des bruits blancs gaussiens indépendants et X_k vérifie maintenant l'équation

$$X_{k+1} = \frac{\sigma^2}{\bar{\theta}_k} [h(X_k) - h(M_k)] + X_k + b(X_k) \varepsilon + \sigma \sqrt{\varepsilon} \tilde{w}_{k+1}. \quad (43)$$

Désignant toujours par S le logarithme de $L_k \Lambda_k$, puisque

$$\begin{aligned} \tilde{w}_i &= \frac{1}{\sigma \sqrt{\varepsilon}} (X_i - M_i) - \frac{1}{\sigma \sqrt{\varepsilon}} (X_{i-1} - M_{i-1}) - \frac{\sqrt{\varepsilon}}{\sigma} [b(X_{i-1}) - b(M_{i-1})] \\ &\quad - \frac{\sigma}{\sqrt{\varepsilon} \bar{\theta}_{i-1}} [h(X_{i-1}) - h(M_{i-1})] + \frac{\bar{\theta}_{i-1}}{\sigma \sqrt{\varepsilon}} [\tilde{y}_i - h(X_{i-1})], \end{aligned}$$

on a

$$\begin{aligned} S &= -\frac{1}{\varepsilon} \sum_{i=1}^k \frac{1}{\bar{\theta}_{i-1}} [h(X_{i-1}) - h(M_{i-1})] [(X_i - M_i) - (X_{i-1} - M_{i-1})] \\ &\quad + \sum_{i=1}^k \frac{1}{\bar{\theta}_{i-1}} [h(X_{i-1}) - h(M_{i-1})] [b(X_{i-1}) - b(M_{i-1})] \\ &\quad + \frac{1}{\varepsilon} \sum_{i=1}^k \left[-\frac{1}{2} h^2(X_{i-1}) + h'(X_{i-1}) h(M_{i-1}) \right] \\ &\quad + \frac{\sigma^2}{2\varepsilon} \sum_{i=1}^k \frac{1}{\bar{\theta}_{i-1}^2} [h(X_{i-1}) - h(M_{i-1})]^2 + \frac{1}{\varepsilon} \sum_{i=1}^k h(M_{i-1}) [\tilde{y}_i - h(M_{i-1})]. \quad (44) \end{aligned}$$

Le dernier terme dans cette égalité est \tilde{Y}^k -adapté et disparaîtra donc dans la normalisation.

Dérivation par rapport à la condition initiale.

Avec les mêmes notations utilisées précédemment on obtient,

- en dérivant l'expression (43),

$$Z_{i+1} = (1 + \frac{\sigma^2}{\bar{\theta}_i} h'(X_i) + b'(X_i) \varepsilon) Z_i,$$

où $\frac{\sigma^2}{\bar{\theta}_i} h'(X_i) > c_h^2$ (voir la preuve du lemme 3.13). On a donc la majoration

$$Z_{nk} \leq (1 + c_h^2 - \|b'\| \varepsilon)^{-(n-k)}.$$

- en dérivant l'expression (44),

$$\begin{aligned} \frac{\partial}{\partial X_0} \log(L_k \Lambda_k) = & -\frac{1}{\varepsilon} \sum_{i=1}^k \frac{1}{\bar{\theta}_{i-1}} (X_i - M_i) h'(X_{i-1}) Z_{i-1} \\ & + \frac{1}{\varepsilon} \sum_{i=1}^k \frac{1}{\bar{\theta}_{i-1}} (X_{i-1} - M_{i-1}) h'(X_{i-1}) Z_{i-1} \\ & + \sum_{i=1}^k \frac{1}{\bar{\theta}_{i-1}} [b(X_{i-1}) - b(M_{i-1})] h'(X_{i-1}) Z_{i-1} \\ & - \frac{1}{\varepsilon} \sum_{i=1}^k [h(X_{i-1}) - h(M_{i-1})] \left\{ \frac{1}{\bar{\theta}_i} [Z_i - Z_{i-1} - b'(X_{i-1}) \varepsilon Z_{i-1}] \right. \\ & \quad \left. - \frac{\sigma^2 - \bar{\theta}_{i-1}^2}{\bar{\theta}_{i-1}^2} h'(X_{i-1}) Z_{i-1} \right\}. \end{aligned}$$

Expression asymptotique pour $X_k - M_k$.

On obtient l'expression suivante:

$$\begin{aligned} X_k - M_k = & \frac{\bar{\theta}_{k-1}}{h'(X_{k-1})} \sum_{i=1}^k (X_{i-1} - M_{i-1}) \left\{ h'(X_{i-1}) \left[\frac{1}{\bar{\theta}_{i-1}} - h'(X_{i-1}) \right] Z_{k-1,i-1} - \frac{h'(X_{i-2})}{\bar{\theta}_{i-2}} Z_{k-1,i-2} \right\} \\ & + \frac{\bar{\theta}_{k-1}}{2 h'(X_{k-1})} \sum_{i=1}^k (X_{i-1} - M_{i-1})^2 h''(\xi_{i-1}^X) h'(X_{i-1}) Z_{k-1,i-1} \\ & + \frac{\bar{\theta}_{k-1}}{h'(X_{k-1})} \varepsilon \sum_{i=1}^k (X_{i-1} - M_{i-1}) b'(\zeta_{i-1}^X) \frac{h'(X_{i-1})}{\bar{\theta}_{i-1}} Z_{k-1,i-1} \\ & - \frac{\bar{\theta}_{k-1}}{h'(X_{k-1})} \varepsilon \frac{\partial}{\partial X_0} \log(L_k \Lambda_k) Z_{k-1,0}. \end{aligned}$$

Mais

• pour $i \geq 2$,

$$\begin{aligned} & h'(X_{i-1}) \left[\frac{1}{\bar{\theta}_{i-1}} - h'(X_{i-1}) \right] Z_{k-1,i-1} - \frac{h'(X_{i-2})}{\bar{\theta}_{i-2}} Z_{k-1,i-2} \\ &= h'(X_{i-1}) \left[\frac{1}{\bar{\theta}_{i-1}} - h'(X_{i-1}) \right] b'(X_{i-2}) \varepsilon Z_{k-1,i-2} + \phi_i Z_{k-1,i-2}, \end{aligned}$$

où

$$\phi_i \triangleq h'(X_{i-1}) \left[\frac{1}{\bar{\theta}_{i-1}} - h'(X_{i-1}) \right] \left[1 + \frac{\sigma^2}{\bar{\theta}_{i-2}} h'(X_{i-2}) \right] - \frac{h'(X_{i-2})}{\bar{\theta}_{i-2}}.$$

Étant donné que $\bar{\theta}_i$ vérifie

$$h'(M_{i-1}) \left[\frac{1}{\bar{\theta}_{i-1}} - h'(M_{i-1}) \right] \left[1 + \frac{\sigma^2}{\bar{\theta}_{i-2}} h'(M_{i-2}) \right] - \frac{h'(M_{i-2})}{\bar{\theta}_{i-2}} = 0$$

et, en utilisant les développements de Taylor jusqu'à l'ordre 2 de la fonction h' , on a

$$\begin{aligned} \phi_i &= \left\{ \frac{1}{\bar{\theta}_{i-1}} + \frac{\sigma^2}{\bar{\theta}_{i-1} \bar{\theta}_{i-2}} h'(X_{i-2}) - [h'(X_{i-1}) + h'(M_{i-1})] \right. \\ &\quad \left. - \frac{\sigma^2}{\bar{\theta}_{i-2}} h'(M_{i-2}) [h'(X_{i-1}) + h'(M_{i-1})] \right\} h''(\bar{\xi}_{i-1}^X) (X_{i-1} - M_{i-1}) \\ &\quad + \left\{ -\frac{1}{\bar{\theta}_{i-2}} + \frac{\sigma^2}{\bar{\theta}_{i-1} \bar{\theta}_{i-2}} h'(M_{i-1}) - h'^2(X_{i-1}) \frac{\sigma^2}{\bar{\theta}_{i-2}} \right\} h''(\bar{\xi}_{i-2}^X) (X_{i-2} - M_{i-2}) \end{aligned}$$

• pour $i = 1$ on a tout simplement

$$\phi_1 \triangleq h'(X_0) \left[\frac{1}{\bar{\theta}_0} - h'(X_0) \right].$$

L'expression de $X_k - M_k$ devient:

$$\begin{aligned} & X_k - M_k \\ &= \frac{\bar{\theta}_{k-1}}{h'(X_{k-1})} \sum_{i=3}^k (X_{i-1} - M_{i-1})(X_{i-2} - M_{i-2}) \\ &\quad \cdot \left\{ -\frac{1}{\bar{\theta}_{i-2}} + \frac{\sigma^2}{\bar{\theta}_{i-1} \bar{\theta}_{i-2}} h'(M_{i-1}) - h'^2(X_{i-1}) \frac{\sigma^2}{\bar{\theta}_{i-2}} \right\} h''(\bar{\xi}_{i-2}^X) Z_{k-1,i-2} \\ &\quad + \frac{\bar{\theta}_{k-1}}{h'(X_{k-1})} \sum_{i=2}^k (X_{i-1} - M_{i-1})^2 \\ &\quad \cdot \left\{ \frac{1}{\bar{\theta}_{i-1}} + \frac{\sigma^2}{\bar{\theta}_{i-1} \bar{\theta}_{i-2}} h'(X_{i-2}) - [h'(X_{i-1}) + h'(M_{i-1})] \left[1 + \frac{\sigma^2}{\bar{\theta}_{i-2}} h'(M_{i-2}) \right] \right\} \end{aligned}$$

$$\begin{aligned}
& \cdot h''(\bar{\xi}_{i-1}^X) Z_{k-1,i-2} \\
& + \frac{\bar{\theta}_{k-1}}{2 h'(X_{k-1})} \sum_{i=1}^k (X_{i-1} - M_{i-1})^2 h''(\bar{\xi}_{i-1}^X) h'(X_{i-1}) Z_{k-1,i-1} \\
& + \frac{\bar{\theta}_{k-1}}{h'(X_{k-1})} \varepsilon \sum_{i=2}^k (X_{i-1} - M_{i-1}) h'(X_{i-1}) \left[\frac{1}{\bar{\theta}_{i-1}} - h'(X_{i-1}) \right] b'(X_{i-2}) Z_{k-1,i-2} \\
& + \frac{\bar{\theta}_{k-1}}{h'(X_{k-1})} \varepsilon \sum_{i=2}^k (X_{i-1} - M_{i-1}) b'(\zeta_{i-1}^X) \frac{h'(X_{i-1})}{\bar{\theta}_{i-1}} Z_{k-1,i-1} \\
& + \frac{\bar{\theta}_{k-1}}{h'(X_{k-1})} (X_1 - M_1)(X_0 - M_0) \left\{ -\frac{1}{\bar{\theta}_0} + \frac{\sigma^2}{\bar{\theta}_1 \bar{\theta}_0} h'(M_1) - h'^2(X_1) \frac{\sigma^2}{\bar{\theta}_0} \right\} h''(\bar{\xi}_0^X) Z_{k-1,0} \\
& + \frac{\bar{\theta}_{k-1}}{h'(X_{k-1})} (X_0 - M_0) \left\{ \frac{1}{\bar{\theta}_0} - h'(X_0) + \varepsilon b'(\zeta_0^X) \frac{1}{\bar{\theta}_0} \right\} h'(X_0) Z_{k-1,0} \\
& - \frac{\bar{\theta}_{k-1}}{h'(X_{k-1})} \varepsilon \frac{\partial}{\partial X_0} \log(L_k \Lambda_k) Z_{k-1,0}.
\end{aligned} \tag{45}$$

Comme dans le cas précédent on cherche des majorations dans L^1 pour les différents termes à droite dans cette expression, à l'exception du dernier.

On a la majoration (voir le lemme 3.13 et la proposition 3.12)

$$\begin{aligned}
& E[|(X_{i-1} - M_{i-1})(X_{i-2} - M_{i-2})|] \\
& \leq E[|1 - \bar{\theta}_k h'(\xi_k^X) + b'(\zeta_k^X) \varepsilon| (X_{i-2} - M_{i-2})^2] \\
& \quad + \sqrt{\varepsilon} E[|\sigma w_{i-1} - \bar{\theta}_{i-2} v_{i-1}| |X_{i-2} - M_{i-2}|] \\
& \leq (1 - A) E[(X_{i-2} - M_{i-2})^2] + C \sqrt{\varepsilon} (E[(X_{i-2} - M_{i-2})^2])^{1/2} \\
& \leq (1 - A)^{2i-3} E[(X_0 - M_0)^2] + C \sqrt{\varepsilon} (1 - A)^{i-2} (E[(X_0 - M_0)^2])^{1/2} + C \varepsilon,
\end{aligned}$$

où la constante A dans le lemme 3.13 peut être choisie telle que $1 - A > (1 + c_h^2 - \|b'\| \varepsilon)^{-1/2}$.

On a alors que:

- le 1^{ier} terme est majoré (dans L^1) par:

$$\begin{aligned}
& C \sum_{i=3}^k \{ C (1 - A)^{2(i-1)} + C \varepsilon + C \sqrt{\varepsilon} (1 - A)^{i-2} \} (1 + c_h^2 - \|b'\| \varepsilon)^{-(k-i+1)} \\
& \leq C \exp\left\{-c \frac{t_k}{\varepsilon}\right\} + C \varepsilon
\end{aligned}$$

- le 2^{me} terme est majoré (dans L^1) par:

$$C \sum_{i=2}^k [(1 - A)^{2(i-1)} + C \varepsilon] (1 + c_h^2 - \|b'\| \varepsilon)^{-(k-i+1)} \leq C \exp\left\{-c \frac{t_k}{\varepsilon}\right\} + C \varepsilon$$

- le 3^{me} terme est majoré (dans L^1) par:

$$C \sum_{i=1}^k [(1-A)^{2(i-1)} + C\varepsilon] (1 + c_h^2 - \|b'\| \varepsilon)^{-(k-i)} \leq C \exp \left\{ -c \frac{t_{k-1}}{\varepsilon} \right\} + C\varepsilon$$

- le 4^{me} terme est majoré (dans L^1) par:

$$C\varepsilon \sum_{i=2}^k [(1-A)^{(i-1)} + C\sqrt{\varepsilon}] (1 + c_h^2 - \|b'\| \varepsilon)^{-(k-i+1)} \leq C\varepsilon \exp \left\{ -c \frac{t_k}{\varepsilon} \right\} + C\varepsilon^{3/2}$$

- le 5^{me} terme est majoré (dans L^1) par:

$$C\varepsilon \exp \left\{ -c \frac{t_{k-1}}{\varepsilon} \right\} + C\varepsilon^{3/2}$$

- les 6^{me} et 7^{me} termes, représentant la contribution de l'erreur à l'instant initial, sont majorés (dans L^1) par:

$$C (1 + c_h^2 - \|b'\| \varepsilon)^{-(k-1)} \leq C \exp \left\{ -c \frac{t_{k-1}}{\varepsilon} \right\}.$$

Finalement, on a que

$$E \left[\left| X_k - M_k + \frac{\bar{\theta}_{k-1}}{h'(X_{k-1})} \varepsilon \frac{\partial}{\partial X_0} \log(L_k \Lambda_k) Z_{k-1,0} \right| \right] \leq C \exp \left\{ -c \frac{t_{k-1}}{\varepsilon} \right\} + C\varepsilon.$$

Comme dans la preuve du théorème 3.3, prenons l'espérance conditionnelle par rapport à \bar{Y}^k dans l'expression de $X_k - M_k$.

Le lemme qui suit complète la démonstration. ■

Lemme 3.15 *On a l'estimation*

$$E \left[\frac{\partial}{\partial X_0} \log(L_k \Lambda_k) Z_{k-1,0} \mid \bar{Y}^k \right] = \mathcal{O}(1).$$

Preuve La preuve de ce lemme suit les mêmes pas de celle du cas $\alpha > 1$. Il suffit de remarquer que Z_{n0} est donné par

$$Z_{n0} = \prod_{i=0}^{n-1} \left(1 + \frac{\sigma^2}{\bar{\theta}_i} h'(X_i) + b'(X_i) \varepsilon \right)^{-1}$$

et sa dérivée par rapport à X_0 est donc

$$\begin{aligned} \frac{\partial}{\partial X_0} Z_{n0} = & - \sum_{i=0}^{n-1} \left(\frac{\sigma^2}{\theta_i} h''(X_i) + b''(X_i) \varepsilon \right) \left(1 + \frac{\sigma^2}{\theta_i} h'(X_i) + b'(X_i) \varepsilon \right)^{-1} \\ & \cdot \prod_{j=i}^{n-1} \left(1 + \frac{\sigma^2}{\theta_j} h'(X_j) + b'(X_j) \varepsilon \right)^{-1}, \end{aligned}$$

d'où la majoration:

$$\left| \frac{\partial}{\partial X_0} Z_{n0} \right| \leq C (1 + c_h^2 - \|b'\| \varepsilon)^{-1} \sum_{i=1}^n (1 + c_h^2 - \|b'\| \varepsilon)^{-1} \leq C.$$

■

Remarque 3.16 Dans le cas d'une fonction h linéaire, on peut vérifier que les termes dans l'expression (45) disparaissent à l'exception des 4^{me}, 5^{me}, 7^{me} et 8^{me} termes. On obtient ainsi une estimation d'ordre $\mathcal{O}(\varepsilon)$ au lieu de l'ordre $\mathcal{O}(1)$ dans le lemme 3.15. On conclut alors que, si h est une fonction linéaire, on a

$$\tilde{X}_k - M_k = \mathcal{O}(\varepsilon^{3/2}),$$

i.e.

$$E[|\tilde{X}_k - M_k|] \leq C \exp\left\{-c \frac{t_k}{\varepsilon}\right\} + C \varepsilon^{3/2}.$$

3.2.3 Estimation de l'erreur de filtrage

De même que dans le paragraphe 3.1.3, on obtient immédiatement une estimation de l'erreur de filtrage à partir de l'estimation de l'erreur $\tilde{X}_k - M_k$.

Corollaire 3.17 On a l'estimation

$$E[|\hat{X}_k - M_k|] \leq C \exp\left\{-c \frac{t_k}{\varepsilon}\right\} + C \varepsilon.$$

3.3 Le cas $\alpha < 1$

Dans le cas où $\Delta t = \varepsilon^\alpha$, avec $\alpha < 1$, il est bien connu que les observations prises dans le passé ne sont pas de grande utilité dans le calcul de l'estimation courante. On fait le raisonnement suivant:

D'après l'équation d'observation on a

$$y_k - h(X_k) = \frac{\varepsilon}{\sqrt{\Delta t}} \bar{w}_k$$

est un bruit blanc gaussien de variance $\varepsilon^{2-\alpha}$, donc il est bien connu que, pour $c_0 \geq \varepsilon^{1-\alpha/2}$, on a la majoration

$$P(|y_k - h(X_k)| > c_0) \leq \frac{2\varepsilon^{1-\alpha/2}}{c_0\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{c_0}{\varepsilon^{1-\alpha/2}}\right)^2\right\}.$$

Supposons que la fonction h est inversible et que h^{-1} est continue et à dérivée bornée.

Prenons

$$M_k = h^{-1}(y_k), \quad (46)$$

i.e. notre filtre approché consiste tout simplement à prendre l'image inverse par h de l'observation.

Alors

$$|X_k - M_k| \leq \|(h^{-1})'\| |h(X_k) - y_k|$$

donc, pour $c_1 \geq \|(h^{-1})'\| \varepsilon^{1-\alpha/2}$, on a que

$$P(|X_k - M_k| > c_1) \leq \frac{2\|(h^{-1})'\| \varepsilon^{1-\alpha/2}}{c_1\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{c_1}{\|(h^{-1})'\| \varepsilon^{1-\alpha/2}}\right)^2\right\}.$$

En termes d'espérances on a tout simplement que

$$E[(X_k - M_k)^2] \leq \|(h^{-1})'\|^2 E[(h(X_k) - y_k)^2] = \|(h^{-1})'\|^2 \frac{\varepsilon^2}{\Delta t}.$$

En un mot, le filtre défini par (46) est déjà un filtre "performant".

Quand au résultat concernant l'estimation de l'erreur par rapport au filtre optimal, il semble difficile de trouver le bon changement de probabilité qui conduit à l'expression asymptotique pour $X_k - M_k$ dont les termes soient facilement estimés.

3.4 Récapitulatif

Le tableau suivant nous donne le filtre approché à prendre selon la valeur de α ainsi que l'estimation sur son efficacité.

<i>filtre approché</i>		<i>estimation de l'erreur</i>	
		$E[(X_k - M_k)^2]$	$E[\hat{X}_k - M_k]$
$\alpha < 1$	$M_k = h^{-1}(y_k)$	$\varepsilon^{2-\alpha}$	
$\alpha = 1$	$M_k = M_{k-1} + b(M_{k-1})\varepsilon + \bar{\theta}_k(y_k - h(M_{k-1}))$ $\bar{\theta}_k = \frac{h'(M_k)[\bar{\theta}_{k-1} + \sigma^2 h'(M_{k-1})]}{h'^2(M_k)[\bar{\theta}_{k-1} + \sigma^2 h'(M_{k-1})] + h'(M_{k-1})}$	$c\varepsilon$	$c\varepsilon$
$\alpha > 1$	$M_k = M_{k-1} + b(M_{k-1})\Delta t + \sigma \frac{\Delta t}{\varepsilon}(y_k - h(M_{k-1}))$	$c\varepsilon$	$c(\varepsilon^{\alpha-1/2} \vee \varepsilon)$

4 Application numérique

Supposons qu'on veut résoudre le problème de filtrage suivant:

À chaque instant $t \in [0, T]$ on voudrait estimer le processus X_t donné par l'équation différentielle stochastique ci dessous, étant connue l'observation Y_t jusqu'à l'instant t

$$\begin{cases} dX_t = -X_t dt + \sigma dw_t^1 \\ dY_t = (X_t^3 + X_t) dt + \varepsilon dw_t^2, \end{cases}$$

où w_t^1 et w_t^2 sont des processus de Wiener indépendants et X_0 est une variable aléatoire gaussienne de loi $\mathcal{N}(\bar{X}_0, p_0)$, avec

- $\sigma = 1.$
- $\varepsilon = 1, 0.1, 0.01, 0.001, 0.0001$
- $\bar{X}_0 = 2.$
- $p_0 = 1.$

La résolution numérique de ce problème peut être envisagée de deux façons distinctes: on peut discrétiser l'équation de Zakai ou utiliser un filtre approché pour calculer une approximation de la solution. Dans notre cas, on commencera par discrétiser le système continu et ainsi obtenir un système discret. Pour ce système discret, on discutera l'application du filtre approché proposé dans la section 3.1 (bien que l'hypothèse (H3') ne soit pas vérifiée dans ce cas). On remarque qu'il s'agit d'un filtre de dimension 1, au sens où on a une seule expression recursive à évaluer. Cet estimateur nous permettra d'obtenir une approximation du filtre optimal facile à calculer. Par ailleurs, on avait démontré des résultats asymptotiques garantissant la qualité du filtre (voir la section 3.1). Les pas de discrétisation à considérer varieront selon la valeur de ε et on regardera l'intervalle de temps $[0, 1]$.

Dans la suite on se propose de vérifier numériquement la performance du filtre approché par rapport à l'estimation donnée par l'équation de Zakai. Le comportement du filtre par rapport à la trajectoire simulée est aussi étudié.

Les 3 figures qui suivent (figure 1, figure 2 et figure 3) représentent l'erreur quadratique moyenne (*eqm*), obtenue sur 100 simulations de trajectoires indépendantes. Pour des raisons évidentes on utilise des échelles logarithmiques sur les deux axes.

On voit immédiatement que, quand ε est petit ($\varepsilon = 0.1, 0.01, \dots$), la courbe logarithmique s'approche d'une droite, i.e.

$$\sqrt{eqm} \simeq 10^{-\text{gamma}} \varepsilon^{\text{pente}}.$$

On peut donc constater que la proposition 3.2 et le théorème 3.1.2, énoncés précédemment, sont vérifiés numériquement.

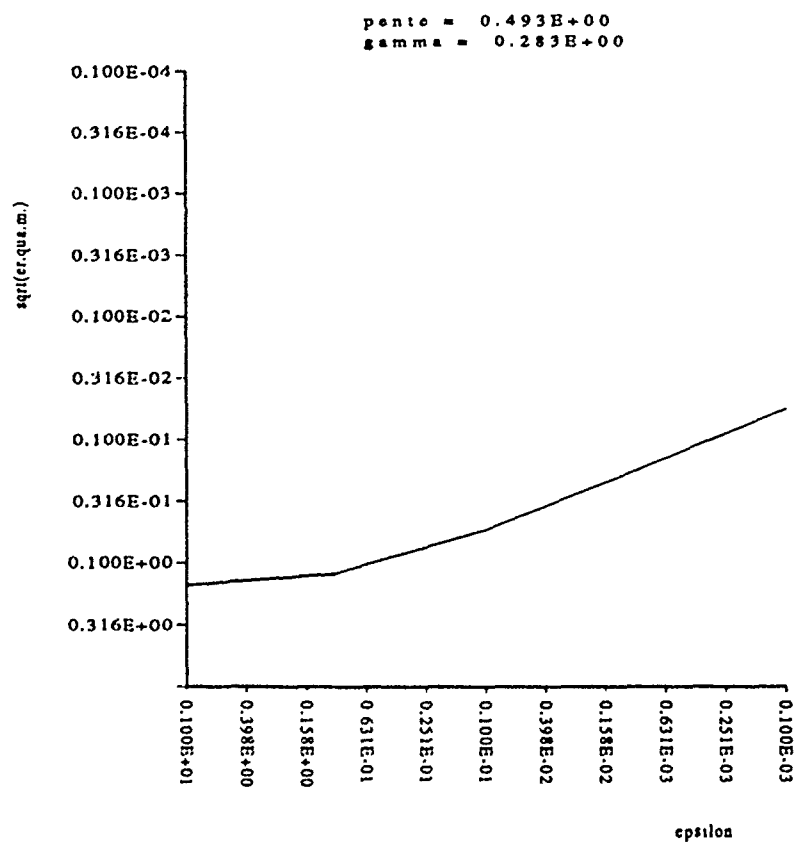


Figure 1: l'erreur quadratique moyenne entre le filtre approché et la trajectoire simulée, en fonction de ϵ .

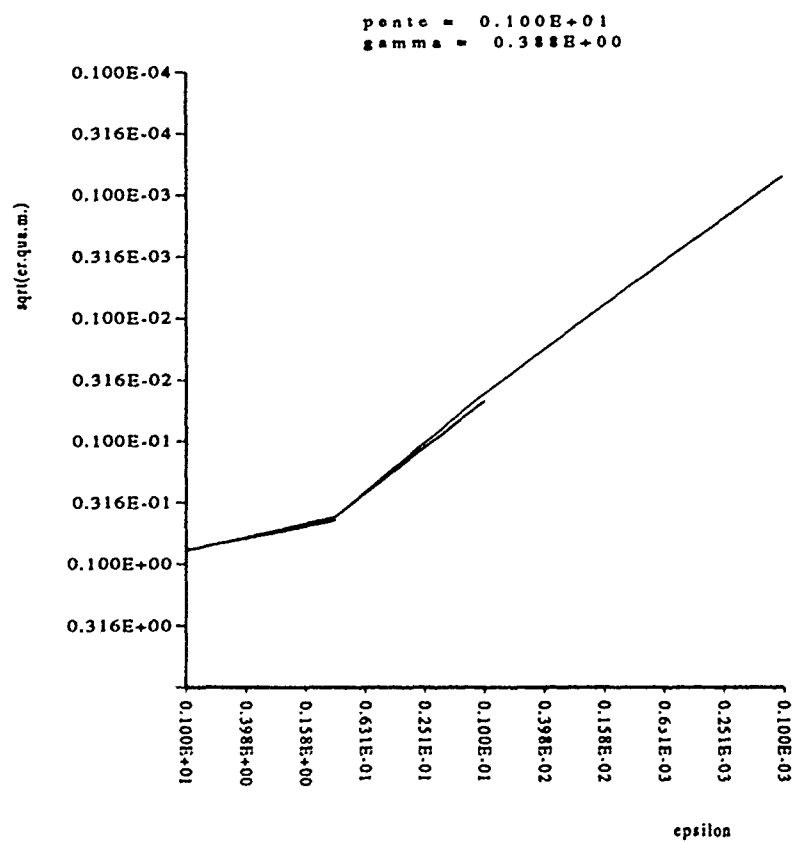


Figure 2: l'erreur quadratique moyenne entre le filtre approché et l'estimation donnée par l'équation de Zakai, en fonction de ϵ .

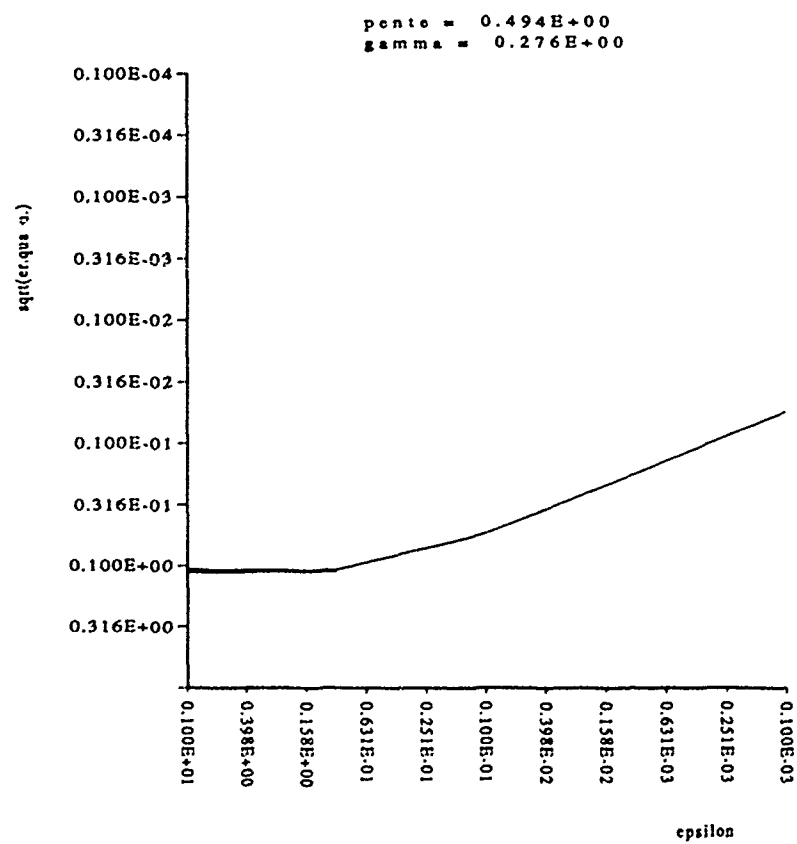


Figure 3: l'erreur quadratique moyenne entre l'estimation donnée par l'équation de Zakai et la trajectoire simulée, en fonction de ε .

Dans la figure 4 on représente, pour $\varepsilon = 0.1$, l'évolution de l'erreur quadratique moyenne au cours du temps. Le pas de discrétisation utilisé a été $\Delta t = 0.01$.

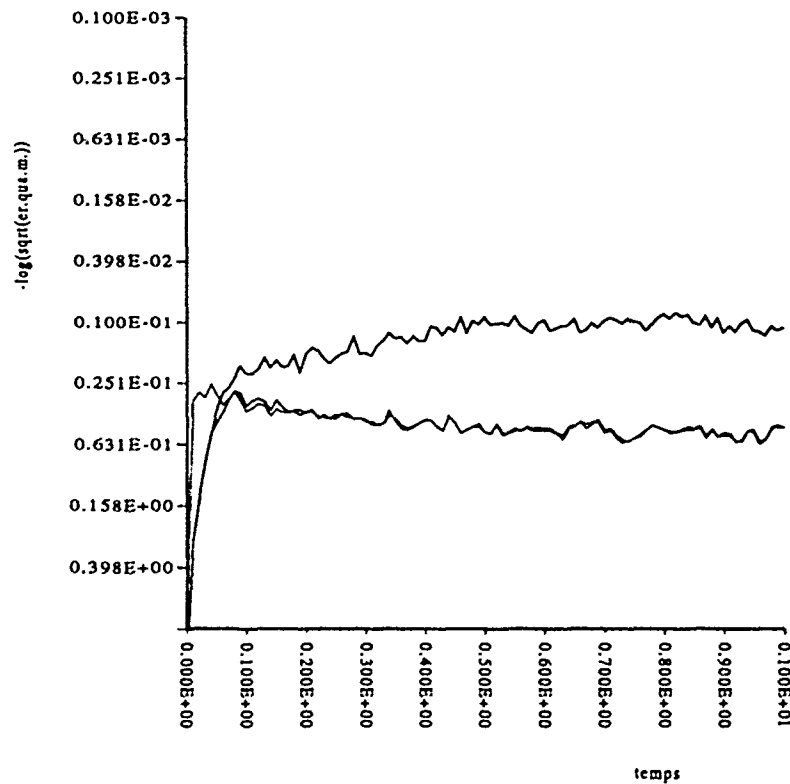


Figure 4: l'erreur quadratique moyenne entre la trajectoire simulée et les estimations et celles-ci entre elles, au cours du temps.

Le trait le plus épais correspond à l'erreur entre les deux filtres et les deux autres traits les erreurs de ces mêmes filtres par rapport à la trajectoire simulée, le trait le plus fin étant celui qui correspond à l'erreur relative au filtre optimal. La ressemblance entre le comportement du filtre approché et du filtre optimal est évidente.

Finalement, dans la dernière figure (figure 5), on représente une seule trajectoire et le filtre approché, ainsi que l'estimation obtenue par résolution numérique de l'équation de Zakai.

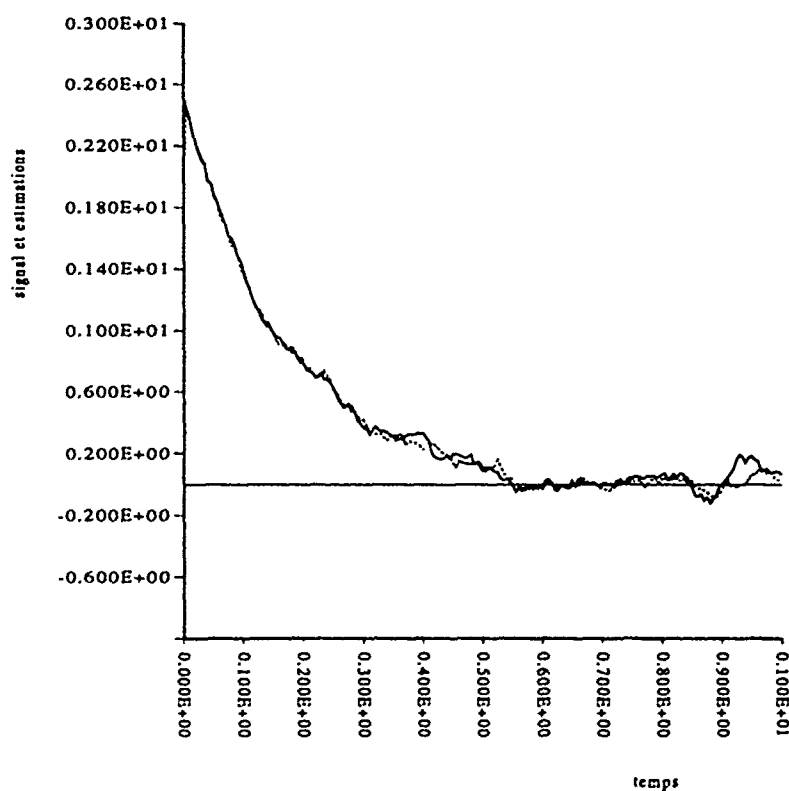


Figure 5: l'évolution des filtres approché et optimal sur une trajectoire simulée.

On voit facilement que les filtres (en pointillé sur la figure) évoluent de manière semblable et qu'il devient difficile de distinguer les deux.

Pour le calcul de la solution de l'équation de Zakai on a utilisé un programme FORTRAN généré à cet effet par le logiciel ZPB, travail en cours de développement au sein du projet MEFISTO, à l'INRIA - centre de Sophia - Antipolis.

Dans cette étude asymptotique de qualité des filtres, deux difficultés importantes se présentent, dans la mise en œuvre des algorithmes. Effectivement, quand $\varepsilon \rightarrow 0$, la densité conditionnelle solution du problème de filtrage s'approche d'une gaussienne ayant un pic "très prononcé", ce qui nous a obligé à utiliser un pas de discrétisation en espace très petit dans l'algorithme de résolution de l'équation de Zakai. Cette difficulté conduira à la recherche d'algorithmes plus performants pour la discrétisation de l'équation de Zakai, lorsque le bruit d'observation est petit.

Les essais numériques ici présentés ont été réalisés sur un mini supercalculateur vectoriel CONVEX C210. Le temps de calcul nécessaire et l'espace de mémoire requise nous ont posé des contraintes pour la résolution numérique du problème de filtrage. D'abord, dans la discrétisation du système continu par rapport au temps, on est obligé de prendre des pas de temps de plus en plus petits, quand ε devient lui aussi petit. Le pas de temps le plus petit considéré dans cet exemple est de l'ordre de 10^{-5} . D'autre part, il y a le fait qu'on vient de citer que le nombre de points nécessaires dans la discrétisation en espace de l'équation de Zakai est très élevé. Dans notre exemple il a été de l'ordre de 2×10^5 pour les valeurs plus petites de ε .

Pour la résolution de ce problème on aurait pu aussi bien utiliser d'autres filtres approchés, par exemple le filtre de Kalman étendu ou encore le filtre de Katzur-Picard, une fois discrétisé (cf. [Milheiro]).

Références

- [Bensoussan] Bensoussan, A. : *On some approximation techniques in non linear filtering*, Proceedings of June 1986 Workshop, Institute for Math. and Its Applications, University of Minnesota.
- [Gelb] Gelb, A. : *Applied Optimal Estimation*, written by the Technical Staff, The Analytical Sciences Corporation, M.I.T. Press, 1974.
- [Jazwinski] Jazwinski, A. H. : *Stochastic Processes and Filtering Theory*, Academic Press, 1970.
- [Kallianpur] Kallianpur, G. : *Stochastic Filtering Theory*, Applications of Mathematics, Tome 13, Springer-Verlag 1975.
- [KBS] Katzur, R. Bobrovsky, B.Z. and Schuss, Z. : *Asymptotic analysis of the optimal filtering problem for one-dimensional diffusions measured in a low noise channel*, SIAM J. Appl. Math. 44 (1984), Part I: 591-604, Part II: 1176-1191.
- [Le Gland] Le Gland, F. : *Estimation de paramètres dans les processus stochastiques, en observation incomplète - Application à un problème de radio-astronomie*, Thèse présentée à l'Université Paris IX-Dauphine le 26 Novembre 1981 pour obtenir le titre de Docteur-Ingenieur.
- [Milheiro] Milheiro de Oliveira, P. : *Filtrage non linéaire avec petit bruit d'observation. Comparaison numérique de 2 filtres*, rapport de stage de D.E.A., Université de Provence, 1987.
- [Picard] Picard, J. : *Nonlinear filtering of one-dimensional diffusions in the case of a high signal-to-noise ratio*. 1986 - SIAM J. Appl. Math. (46), 6, 1098-1125.
- [Picard] Picard, J. : *Asymptotic study of estimation problems with small observation noise*, dans Stochastic Modelling and Filtering (Rome 1984), Lect. Notes in Control and Inf. Sci. 91, Springer, 1987.
- [Picard] Picard, J. : *Filtrage de diffusions vectorielles faiblement bruitées*, Proc. 7th Int. Conf. on Analysis and Optimization of Systems (Antibes 1986), Lect. N. in Control and Information Sciences, 83, Springer 1986.

[Picard]

Picard, J. : *Nonlinear filtering and smoothing with high signal-to-noise ratio*, Stochastic Processes in Physics and Engineering (Bielefeld 1986), Reidel 1988.

A Version discrète du théorème de Girsanov

Considérons une filtration $(\mathcal{F}_k)_k$, $k \in \{0, 1, \dots, K\}$.

Soient:

- $\{w_k\}$ un \mathcal{F}_k -bruit blanc gaussien pour la probabilité P .
- $\{\varphi_k\}$ un processus \mathcal{F}_{k-1} -mesurable (i.e. prévisible) tel que:

$$\sum_{i=0}^K |\varphi_i|^2 < \infty \quad (P \text{ p.s.})$$

- $\{Z_k\}$ le processus défini par:

$$\begin{aligned} Z_k &= \exp\left\{\sum_{i=1}^k \varphi_i w_i - \frac{1}{2} \sum_{i=1}^k |\varphi_i|^2\right\}, \quad k \geq 1 \\ Z_0 &= 1. \end{aligned}$$

Alors,

1. $\{Z_k\}$ est une martingale discrète.

Preuve

$$\begin{aligned} E[Z_k | \mathcal{F}_{k-1}] &= E\left[Z_{k-1} \frac{Z_k}{Z_{k-1}} | \mathcal{F}_{k-1}\right] \\ &= Z_{k-1} E\left[\exp\left\{\varphi_k w_k - \frac{1}{2} |\varphi_k|^2\right\} | \mathcal{F}_{k-1}\right], \\ &\quad \text{puisque } Z_{k-1} \text{ est } \mathcal{F}_{k-1}\text{-mesurable.} \\ &= Z_{k-1}, \\ &\quad \text{puisque } \{w_k\} \text{ est un bruit blanc gaussien et } \varphi_k \text{ est } \mathcal{F}_{k-1}\text{-mesurable.} \end{aligned}$$

et Z_k est \mathcal{F}_k intégrable. ■

On considère la probabilité \tilde{P} définie par:

$$d\tilde{P}(\omega) = Z_K(\omega) dP(\omega). \quad (47)$$

Remarque A.1

Du fait que $\{Z_k\}$ est une martingale, vient que:

$$\frac{d\bar{P}}{dP} \Big|_{\mathcal{F}_k} = Z_k .$$

2. (Version discrète du Théorème de Girsanov)

Pour la probabilité \bar{P} définie dans (47), le processus $\{\bar{w}_k\}$ défini par:

$$\bar{w}_k = w_k - \varphi_k$$

est un bruit blanc gaussien .

Preuve

On veut démontrer que:

$$\bar{E}[\exp(\lambda \bar{w}_k) | \mathcal{F}_{k-1}] = \exp\left(\frac{\lambda^2}{2}\right), \quad \forall \lambda \in \mathbb{R}$$

i.e.

$$\bar{E}[\exp(\lambda \bar{w}_k - \frac{\lambda^2}{2}) | \mathcal{F}_{k-1}] = 1, \quad \forall \lambda \in \mathbb{R} .$$

Or,

$$\begin{aligned} \bar{E}[\exp(\lambda \bar{w}_k - \frac{\lambda^2}{2}) | \mathcal{F}_{k-1}] &= \frac{E[\exp(\lambda \bar{w}_k - \frac{\lambda^2}{2}) Z_k | \mathcal{F}_{k-1}]}{E[Z_k | \mathcal{F}_{k-1}]} \quad (\text{formule de Bayes}) \\ &= \frac{E[\exp(\lambda w_k - \lambda \varphi_k - \frac{\lambda^2}{2}) Z_k | \mathcal{F}_{k-1}]}{Z_{k-1}}, \end{aligned}$$

puisque Z_k est une martingale et d'après la remarque ci dessus.

$$\begin{aligned} &= E[\exp(\lambda w_k - \lambda \varphi_k - \frac{\lambda^2}{2}) \exp(\varphi_k w_k - \frac{1}{2} \varphi_k^2) | \mathcal{F}_{k-1}] \\ &= E\{\exp[(\lambda + \varphi_k) w_k - \frac{1}{2}(\lambda + \varphi_k)^2] | \mathcal{F}_{k-1}\} \\ &= 1, \text{ de même qu'en 1. } \end{aligned}$$

■

B Estimation du premier terme dans l'expression (30)

Dans le contexte du paragraphe sur l'expression asymptotique de $X_k - M_k$, dans la section 3.1.2, soit

$$S_k \triangleq \sum_{i=2}^k (X_{i-1} - M_{i-1})(X_{i-1} - X_{i-2}) h''(X_{i-2}) Z_{k-1,i-2}$$

(S_k est \mathcal{F}_{k-1} - mesurable)

$$S_0 \triangleq 0.$$

Notre but est de majorer $E|S_k|$.

On établit le résultat suivant:

Lemme B.1 *On a l'estimation*

$$E(S_{k+1}^2) \leq C \varepsilon \exp\left\{-c \frac{t_k}{\varepsilon}\right\} + C \varepsilon^2.$$

Preuve

Notation:

Dans la suite, P_{k-1} désignera la partie \mathcal{F}_{k-1} -mesurable de $X_k - M_k$ i.e.

$$P_{k-1} = (X_{k-1} - M_{k-1}) + \Delta t [b(X_{k-1}) - b(M_{k-1})]$$

$$- \frac{\sigma \Delta t}{\varepsilon} [h(X_{k-1}) - h(M_{k-1})]$$

$$X_k - M_k = P_{k-1} + \sigma \sqrt{\Delta t} (w_k - v_k).$$

D'après la définition de S_k , ci dessus, on a que

$$S_{k+1} = S_k + (X_k - M_k)(X_k - X_{k-1}) h''(X_{k-1}) Z_{k,k-1}.$$

- On commence par estimer le facteur de $Z_{k,k-1}$ dans cette expression:

$$E[(X_k - M_k)^2 (X_k - X_{k-1})^2 h''^2(X_{k-1})]$$

$$= \Delta t^2 E[(X_k - M_k)^2 b^2(X_{k-1}) h''^2(X_{k-1})]$$

$$+ \sigma^2 \Delta t E[(X_k - M_k)^2 w_k^2(X_{k-1}) h''^2(X_{k-1})]$$

$$\begin{aligned}
& + 2\sigma \Delta t^{3/2} E[(X_k - M_k)^2 b(X_{k-1}) w_k h''(X_{k-1})] \\
& \leq C \Delta t E(P_{k-1}^2) + C \Delta t^2 [1 + C E[(X_k - M_k)^2] + C E[|P_{k-1}|]] \\
& \leq C \Delta t [(1 - c_h \frac{\Delta t}{\varepsilon} + \|b'\| \Delta t)^{2k} + C \varepsilon],
\end{aligned}$$

puisque

$$E(P_{k-1}^2) \leq E[(X_k - M_k)^2] \leq (1 - c_h \frac{\Delta t}{\varepsilon} + \|b'\| \Delta t)^{2k} + C \varepsilon.$$

- On estime le facteur $S_k (X_k - M_k)(X_k - X_{k-1}) h''(X_{k-1})$:

$$\begin{aligned}
& E[S_k (X_k - M_k)(X_k - X_{k-1}) h''(X_{k-1})] \\
& = \Delta t E[b(X_{k-1}) h''(X_{k-1}) S_k (X_k - M_k)] \\
& \quad + \sigma \sqrt{\Delta t} E[b(X_{k-1}) h''(X_{k-1}) S_k (X_k - M_k) w_k] \\
& \leq C \Delta t (E(S_k^2))^{1/2} (E[(X_k - M_k)^2])^{1/2} + C \Delta t (E(S_k^2))^{1/2} \\
& \leq \frac{c_h}{2} \frac{\Delta t}{\varepsilon} E(S_k^2) + C \Delta t [(1 - c_h \frac{\Delta t}{\varepsilon} + \|b'\| \Delta t)^{2k} + C \varepsilon].
\end{aligned}$$

On obtient enfin la majoration:

$$\begin{aligned}
& E(S_{k+1}^2) \\
& \leq (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-2} \{ (1 + c_h \frac{\Delta t}{\varepsilon}) E(S_k^2) \\
& \quad + C \Delta t [(1 - c_h \frac{\Delta t}{\varepsilon} + \|b'\| \Delta t)^{2k} + C \varepsilon] \} \\
& \leq C \Delta t (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-2} \sum_{i=0}^{k-1} [(1 + c_h \frac{\Delta t}{\varepsilon}) - \|b'\| \Delta t]^{-2} (1 + c_h \frac{\Delta t}{\varepsilon})^i \\
& \quad \cdot [(1 - c_h \frac{\Delta t}{\varepsilon} + \|b'\| \Delta t)^{2(k-i)} + C \varepsilon] \\
& \leq C \Delta t (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-2} \left[\frac{1 + c_h \Delta t/\varepsilon}{(1 + c_h \Delta t/\varepsilon - \|b'\| \Delta t)^2} \right]^k \sum_{i=1}^k \left[\frac{[1 - (c_h \Delta t/\varepsilon - \|b'\| \Delta t)^2]^2}{1 + c_h \Delta t/\varepsilon} \right] \\
& \quad + C \varepsilon \Delta t (1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^{-2} \sum_{i=0}^{k-1} \left[\frac{1 + c_h \Delta t/\varepsilon}{(1 + c_h \Delta t/\varepsilon - \|b'\| \Delta t)^2} \right]^i.
\end{aligned}$$

Or

$$(1 + c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^2 - (1 + c_h \frac{\Delta t}{\varepsilon}) \geq c \frac{\Delta t}{\varepsilon} > 0$$

et

$$(1 + c_h \frac{\Delta t}{\varepsilon}) - [1 - (c_h \frac{\Delta t}{\varepsilon} - \|b'\| \Delta t)^2]^2 \geq c_h \frac{\Delta t}{\varepsilon} > 0$$

donc

$$E(S_{k+1}^2) \leq C \varepsilon \exp\{-c \frac{t_k}{\varepsilon}\} + C \varepsilon^2 .$$

■

2

Filtrage linéaire par morceaux
d'un système en temps discret
avec petit bruit d'observation

FILTRAGE LINEAIRE PAR MORCEAUX D'UN SYSTEME EN TEMPS DISCRET AVEC PETIT BRUIT D'OBSERVATION

DISCRETE TIME PIECEWISE LINEAR FILTERING WITH SMALL OBSERVATION NOISE

Résumé

Nous nous intéressons à un problème de filtrage linéaire par morceaux en temps discret avec petit bruit d'observation. Nous présentons et comparons plusieurs tests permettant de déterminer les intervalles de linéarité de la fonction d'observation, notamment dans le cas où elle est symétrique. Sur chacun de ces intervalles, nous approchons le filtre optimal par le filtre de Kalman-Bucy correspondant. Comme dans [10], nous approchons des processus discrets par des diffusions pour estimer les probabilités d'erreur et les temps moyens pour prendre une décision.

Abstract

We are interested in a piecewise linear discret time filtering problem with small observation noise. We present and compare different tests which enable us to compute the intervals of linearity of the observation function, specially when it is symmetric. Over such an interval, we can then approximate the optimal filter by the corresponding Kalman-Bucy filter. As in [10], we approximate discret time processes by diffusions, in order to estimate de probabilities of error and the expected times for taking a decision.

1 Introduction

On s'intéresse au problème de filtrage non linéaire suivant:

$$\begin{cases} dX_t = b(X_t) dt + \sigma(X_t) dU_t \\ dY_t = h(X_t) dt + \varepsilon dV_t, \quad Y_0 = 0, \end{cases} \quad (1)$$

où $\{X_t, t \geq 0\}$ est le processus à valeurs dans \mathbb{R} non observé à estimer à l'instant τ au vu des observations jusqu'à l'instant τ du processus unidimensionnel $\{Y_t, t \geq 0\}$, ε un paramètre "petit" et $\{U_t, t \geq 0\}$ et $\{V_t, t \geq 0\}$ sont des processus de Wiener standards indépendants, à valeurs dans \mathbb{R} .

Il est bien connu que le problème est de dimension infinie, au sens où, pour le résoudre, on a à déterminer la solution d'une équation aux dérivées partielles, par exemple l'équation de Zakai. Si la fonction h est monotone, sous certaines hypothèses générales de "régularité", le filtre de Kalman étendu, entre autres, est une "bonne" approximation du filtre optimal (cf. [Picard], [KBS], [Bensoussan], [Ji] et [Milheiro]). Le problème avec h non monotone, sous une certaine "hypothèse de détectabilité" a été traité dans [Fleming-Pardoux].

Dans le cadre du filtrage linéaire par morceaux, on suppose que

- $b(x) = B_- x 1_{\{x < 0\}} + B_+ x 1_{\{x \geq 0\}}$
- $\sigma(x) = \sigma_- 1_{\{x < 0\}} + \sigma_+ 1_{\{x \geq 0\}}$
- $h(x) = H_- x 1_{\{x < 0\}} + H_+ x 1_{\{x \geq 0\}}$.

Si $H_+ H_- > 0$, i.e. la fonction h est monotone, pour $\varepsilon = 0$, X_t peut être parfaitement connu et pour ε "petit", le problème n'offre pas de grandes difficultés.

Dans le cas $H_+ H_- < 0$, i.e. h non monotone, bien que $h(X_t)$ puisse être estimé de façon précise, il n'est pas immédiat qu'il en soit de même pour X_t . Le filtre de Kalman étendu est en général inefficace, le problème étant de déterminer le signe de $\{X_t, t \geq 0\}$. La détermination du signe n'étant possible que dans le cas où la variance conditionnelle est petite ⁽¹⁾, on introduit l'hypothèse de "détectabilité" suivante notée (HD) :

$$(HD) \quad \begin{cases} H_-^2 \sigma_-^2 \neq H_+^2 \sigma_+^2 & (HD1) \\ \text{ou} \\ H_-^2 \sigma_-^2 = H_+^2 \sigma_+^2 \text{ et } B_+ \neq B_- & (HD2). \end{cases}$$

¹Contre-exemple: $b = 0, h(x) = |x|, \sigma = 1$.

Sous (HD1), dans [FJP], un filtre approché a été proposé et dans [Roubaud], il est montré que sous (HD2) un filtre du même type peut être utilisé comme approximation du filtre optimal.

L'idée est de construire un filtre approché à partir de deux filtres de Kalman-Bucy, associés respectivement aux deux problèmes de filtrage linéaire suivants:

$$\begin{cases} dX_t = B_+ X_t dt + \sigma_+ dU_t \\ dY_t = H_+ X_t dt + \varepsilon dU_t \end{cases} \quad (2)$$

et

$$\begin{cases} dX_t = B_- X_t dt + \sigma_- dU_t \\ dY_t = H_- X_t dt + \varepsilon dU_t. \end{cases} \quad (3)$$

On calcule ces deux filtres "en parallèle" et on utilise deux tests.

- Un premier test permet de détecter des intervalles de temps durant lesquels la trajectoire de $\{X_t\}_t$ ne passe pas par 0.
- Un second test permet de décider si $X_t < 0$ ou $X_t > 0$ sur ces intervalles, sous l'hypothèse (HD1) ou (HD2).

Sur chaque intervalle de monotonie, nous approchons alors le filtre optimal par le filtre de Kalman-Bucy correspondant.

Nous considérons l'intervalle de temps fini $[0, T]$.

Dans ce rapport nous nous intéressons à la résolution numérique du problème linéaire par morceaux dans la situation où la fonction d'observation h est non monotone. Nous commençons par discrétiser le système continu (1) par un schéma classique de discretisation en temps avec pas $\Delta t = \varepsilon$. Le processus $\{X_{k\Delta t}\}_k$ est approximé par $\{x_k\}_k$, $\{U_{(k+1)\Delta t} - U_{k\Delta t}\}_k$ par $\{\sqrt{\Delta t} u_k\}_k$, $\{V_{(k+1)\Delta t} - V_{k\Delta t}\}_k$ par $\{\sqrt{\Delta t} v_k\}_k$, $\{(\Delta t)^{-1} (Y_{(k+1)\Delta t} - Y_{k\Delta t})\}_k$ par $\{y_k\}_k$.

Nous obtenons le modèle discret suivant:

$$\begin{cases} x_{k+1} = x_k + \varepsilon b(x_k) + \sqrt{\varepsilon} \sigma(x_k) u_k \\ y_k = h(x_k) + \sqrt{\varepsilon} v_k, \end{cases} \quad (4)$$

avec les fonctions b , σ et h définies comme précédemment. On émet les hypothèses suivantes:

- (H1) $\{u_k\}_k$ et $\{v_k\}_k$ sont des bruits blancs gaussiens standards et indépendants,
- (H2) x_0 est une variable aléatoire réelle telle que $E(\exp \{c_0 x_0^2\}) < +\infty$, pour un certain $c_0 > 0$,

(H3) $H_- H_+ < 0$ et $\sigma_- \sigma_+ \neq 0$.

Sans réduire la généralité du problème, on suppose que $h(x) \geq 0 \forall x \in \mathbb{R}$, i.e. on suppose que

(H3') $H_- < 0$, $H_+ > 0$ et $\sigma_- \sigma_+ \neq 0$.

Nous adaptons l'étude faite dans le cas continu au cas discret. Notre but est de mettre en œuvre les tests permettant de séparer les intervalles de positivité et de négativité de $\{x_k\}_k$ et de comparer leur performance. Nous nous intéressons principalement à la situation où l'hypothèse (HD2) est vérifiée, le problème sous l'hypothèse (HD1) ayant été traité dans [FJSZ].

L'étude de l'évolution de la loi conditionnelle, obtenue de façon approchée par résolution de l'équation de Zakai discretisée (cf. [Le Gland]), nous aidera à interpréter les résultats obtenus.

Ce rapport est organisé comme il suit:

Dans le paragraphe 2, nous présenterons les deux filtres de Kalman-Bucy (7) et (8) utilisés et nous formulerons des remarques préliminaires. Dans le paragraphe 3 nous nous intéresserons à la détection des passages du signal par 0. En analogie avec le cas continu deux tests seront proposés, l'un basé sur les accroissements des observations (cf. [FJP]) et l'autre sur la sortie des filtres de Kalman-Bucy (cf. [Roubaud]). Sachant que sur un intervalle de temps $[a, b]$ le signal ne passe pas par 0, avec une probabilité donnée, nous considérons un autre type de tests pour décider du signe de $\{x_k\}_k$ sur $[a, b]$. Sous l'hypothèse (HD1), des études ont été faites par [FJSZ]. Ces auteurs ont mis en œuvre deux tests différents: l'un basé sur la variation quadratique, l'autre du type rapport de vraisemblance basé sur les sorties des filtres de Kalman-Bucy. Nous résumerons brièvement ces procédures dans le paragraphe 4. Sous l'hypothèse (HD2), nous présenterons dans le paragraphe 5 un test de rapport de vraisemblance basé sur les accroissements des observations et adapterons le test sur les sorties des filtres de Kalman-Bucy du paragraphe précédent au cas traité.

La mise en œuvre des tests de détection et de décision ainsi décrits nécessite la détermination de formules explicites pour les bornes. Bien qu'une démonstration rigoureuse n'ait pu être faite, les formules proposées sont justifiées de manière heuristique. Dans le cas où $B_- \neq B_+$, nous émettrons l'hypothèse supplémentaire:

(H4) $B_+ < 0$ et $B_- < 0$.

Cette hypothèse n'est pas trop restrictive pour notre propos. En effet, suivant le signe des coefficients de dérive, nous avons deux comportements distincts du

signif. Si B_+ et B_- sont négatifs, il devient stationnaire dès que l'intervalle de monotonie est assez grand. Sinon, il peut "fuir" vers l'infini. En fait, nous étudions la situation la plus délicate au sens où, la probabilité pour que le processus $\{x_k\}$ passe par zéro est importante. Cependant, si les valeurs $|B_-|$ et $|B_+|$ sont grandes, $\{x_k\}$ aura tendance à changer rapidement de signe et les tests n'auront probablement pas suffisamment de temps pour décider.

Le paragraphe 6 est consacré à l'application numérique de ces méthodes. Des critères de comparaison entre les différents tests seront proposés et les résultats des applications à divers exemples seront présentés. D'autre part, la solution de l'équation de Zakai sera utilisée pour la justification du comportement des filtres approchés.

Pour conclure, on présente dans le paragraphe 7 quelques réflexions sur le comportement de ces filtres et on discute la performance des différents tests étudiés.

Notation 1.1 *Etant donné un processus $\{G_k\}$ on écrira*

$$G_k = \mathcal{O}(\varepsilon^q),$$

où $q \in \mathbb{R}_+$, pour signifier qu'il existe $c_1, c_2, c_3 > 0$ tels que

$$E[G_k^2] \leq c_1 \exp\{-c_2 k \varepsilon\} + c_3 \varepsilon^{2q}, \forall k \in \mathcal{N}.$$

Notation 1.2 *On notera c ou C des constantes indépendantes de ε , sans se soucier de leur valeur. Ces constantes pourront donc prendre des valeurs différentes d'une ligne à l'autre.*

2 Deux filtres de Kalman-Bucy en parallèle

Soit \mathcal{Y}_0^k , la tribu des observations jusqu'à l'instant k , $\mathcal{Y}_0^k = \sigma\{y_0, y_1, \dots, y_k\}$. La solution du problème de filtrage associé au système (1) étant donnée par la loi conditionnelle de x_k sachant \mathcal{Y}_0^k , nous commentons brièvement son comportement (cf. [FJSZ]). Sous l'hypothèse de "défectabilité" (HD), la variance de cette loi est "petite", sa densité se concentrant autour de deux maxima locaux, l'un négatif et l'autre positif. Ce phénomène se justifie du fait qu'on a soit $x_k - y_k/H_- = \sqrt{\varepsilon}/H_- v_k$ soit $x_k - y_k/H_+ = \sqrt{\varepsilon}/H_+ v_k$, donc la densité conditionnelle est "petite" excepté sur deux intervalles d'amplitude d'ordre $\mathcal{O}(\sqrt{\varepsilon})$, autour de y_k/H_- et de y_k/H_+ . En emettant l'hypothèse de défectabilité (HD) nous nous plaçons dans la situation où lorsque $\{x_k\}$ prend ses valeurs loin de 0 durant un certain intervalle de temps, un des pics tend à disparaître, la loi conditionnelle s'approchant d'une loi gaussienne. Il apparaît alors légitime d'approcher le filtre optimal par un filtre de Kalman-Bucy. Intuitivement on s'attend à ce que, sous l'hypothèse ($HD2$), le temps nécessaire pour faire "disparaître" l'un des deux pics soit plus long que sous l'hypothèse ($HD1$). Cette idée sera illustrée dans le paragraphe 6 (voir les figures 3 à 6 et 12 à 14).

Soient (\hat{x}_k^+, Q_k^+) et (\hat{x}_k^-, Q_k^-) les filtres de Kalman-Bucy associés respectivement aux systèmes linéaires

$$\begin{cases} x_{k+1} &= (1 + B_+ \varepsilon) x_k + \sqrt{\varepsilon} \sigma_+ u_k \\ y_k &= H_+ x_k + \sqrt{\varepsilon} v_k, y_0 = 0 \end{cases} \quad (5)$$

et

$$\begin{cases} x_{k+1} &= (1 + B_- \varepsilon) x_k + \sqrt{\varepsilon} \sigma_- u_k \\ y_k &= H_- x_k + \sqrt{\varepsilon} v_k, y_0 = 0, \end{cases} \quad (6)$$

avec conditions initiales gaussiennes

- $\hat{x}_0^+ = \hat{x}_0^- = E(x_0)$
- $Q_0^+ = \varepsilon \frac{\sigma_+^2}{1 + H_+^2 \sigma_+^2}, Q_0^- = \varepsilon \frac{\sigma_-^2}{1 + H_-^2 \sigma_-^2}.$

Ils sont donnés par les équations suivantes:

$$\begin{cases} \hat{x}_{k+1}^+ &= (1 + \varepsilon B_+) \hat{x}_k^+ + \frac{1}{\varepsilon} H_+ Q_k^+ (y_{k+1} - H_+ (1 + \varepsilon B_+) \hat{x}_k^+) \\ Q_{k+1}^+ &= \varepsilon \frac{(1 + \varepsilon B_+)^2 Q_k^+ + \sigma_+^2 \varepsilon}{(1 + H_+^2 \sigma_+^2) \varepsilon + H_+^2 (1 + \varepsilon B_+)^2 Q_k^+} \end{cases} \quad (7)$$

et

$$\begin{cases} \hat{x}_{k+1}^- = (1 + \varepsilon B_-) \hat{x}_k^- + \frac{1}{\varepsilon} H_- Q_k^- (y_{k+1} - H_- (1 + \varepsilon B_-) \hat{x}_k^-) \\ Q_{k+1}^- = \varepsilon \frac{(1 + \varepsilon B_-)^2 Q_k^- + \sigma_-^2 \varepsilon}{(1 + H_-^2 \sigma_-^2) \varepsilon + H_-^2 (1 + \varepsilon B_-)^2 Q_k^-} \end{cases} \quad (8)$$

Remarque 2.1 Le fait d'avoir considéré une condition initiale gaussienne pour les systèmes (5) et (6) n'est pas restrictif pour le problème de départ. En effet, puisque les filtres étudiés se caractérisent par le fait d'avoir une "mémoire courte", l'influence de la condition initiale tend à disparaître rapidement.

On note les gains à l'instant k :

$$K_k^+ = \frac{1}{\varepsilon} H_+ Q_k^+ \quad \text{et} \quad K_k^- = \frac{1}{\varepsilon} H_- Q_k^-.$$

Remarque 2.2 Dans une situation stationnaire, les expressions des variances des lois conditionnelles, Q^+ et Q^- , sont facilement calculées. Elles sont d'ordre $\mathcal{O}(\varepsilon)$ et donc les gains stationnaires K_+ et K_- sont d'ordre $\mathcal{O}(1)$ (cf. [Milheiro]).

On introduit les processus dits "d'innovation"

$$\nu_{k+1}^+ = y_{k+1} - H_+ (1 + \varepsilon B_+) \hat{x}_k^+$$

et

$$\nu_{k+1}^- = y_{k+1} - H_- (1 + \varepsilon B_-) \hat{x}_k^-.$$

Les processus ainsi définis sont considérés approximativement comme des "bruits blancs" au sens où leurs variances sont asymptotiquement d'ordre $\mathcal{O}(\varepsilon)$, tandis que les corrélations $E(\nu_{k+1}^+ \nu_k^+)$ et $E(\nu_{k+1}^- \nu_k^-)$ sont d'ordre $\mathcal{O}(\varepsilon^2)$.

On notera $\varepsilon \vartheta_k^+$ et $\varepsilon \vartheta_k^-$ leurs variances respectives et $\varepsilon \vartheta^+$ et $\varepsilon \vartheta^-$ les variances stationnaires correspondantes:

$$\begin{aligned} \varepsilon \vartheta^+ &= \varepsilon (1 + \sigma_+^2 H_+^2) + H_+^2 (1 + \varepsilon B_+)^2 Q^+, \\ \varepsilon \vartheta^- &= \varepsilon (1 + \sigma_-^2 H_-^2) + H_-^2 (1 + \varepsilon B_-)^2 Q^-. \end{aligned}$$

Dans la suite, on étudiera la procédure de test permettant de séparer les intervalles de monotonie de la fonction h et de décider du signe de x_k . Sur chacun de ces intervalles, il sera alors possible d'approcher le filtre optimal par le filtre de Kalman-Bucy correspondant.

3 Tests de detection des passages de $\{x_k\}$ à zéro

Nous devons premièrement déterminer, sur l'intervalle de temps $[0, T]$, les intervalles sur lesquels $\{x_k\}$ est différent de zéro, avec une probabilité donnée, proche de 1. Nous présentons deux tests de détection, le premier s'appliquant aux observations y_k , le second à la sortie de l'un des filtres de Kalman-Bucy.

Soit l'intervalle de temps $[a, b]$, avec $0 < a < b < T$. On pose $m = [(b - a)/\varepsilon]$ et $i_0 = [a/\varepsilon]$. On considère les deux événements suivants:

$$A_- = \{x_k < 0; k = i_0, i_0 + 1, \dots, i_0 + m\},$$

$$A_+ = \{x_k > 0; k = i_0, i_0 + 1, \dots, i_0 + m\}.$$

On remarque que

$$(A_- \cup A_+)^c = \{x_k x_{k+1} < 0 \text{ pour un certain } k, i_0 \leq k \leq i_0 + m\}.$$

3.1 Test basé sur les observations y_k

Etant donné que " $h(x) = 0$ ssi $x = 0$ ", $h(x_k)$ "petit" implique par continuité de l'application h , x_k proche de zéro. Cependant $h(x_k)$ n'est pas observé mais, pour ε "petit", il est approché par y_k .

Soit $c_{obs} > 0$ une constante à déterminer.

On définit l'événement test suivant:

$$C = \{|y_k| \geq c_{obs}\}.$$

On montre de façon similaire à la version en temps continu (cf. [FJP][proposition 3.1]) la proposition suivante (cf. [FJSZ][proposition 2.1]):

Proposition 3.1 Soient $c_{obs} > 0$ et $\varepsilon_0 > 0$ donnés. Il existe $\beta > 0$ tel que $\forall \varepsilon \in (0, \varepsilon_0]$ on a

$$P(\{(A_- \cup A_+)^c | C\}) \leq e^{-\beta/\varepsilon}.$$

Détermination de la constante c_{obs} :

Nous utilisons, pour la détermination de la constante c_{obs} , un raisonnement similaire à celui proposé dans [FJSZ], sauf que dans notre cas le coefficient de dérive n'est plus supposé constant mais constant par morceaux. Intuitivement le choix de c_{obs} doit être un compromis entre "conserver des intervalles de monotonie suffisamment grands" et "réduire la probabilité d'erreur du test".

On cherche à ce que la probabilité

$$P(\{x_k x_{k+1} < 0\} \mid \{|y_k| \geq c_{obs}, |y_{k+1}| \geq c_{obs}\})$$

soit faible.

D'après l'hypothèse (H3) on obtient:

$$P(\{y_k < -c_{obs}\}) \leq P(\{v_k < -c_{obs}/\sqrt{\varepsilon}\}) \leq \frac{1}{2} e^{-c_{obs}^2/(2\varepsilon)}.$$

Nous nous permettons donc de considérer uniquement le conditionnement par rapport à l'événement $\{y_k \geq c_{obs}, y_{k+1} \geq c_{obs}\}$. De plus, on a:

$$\{x_k x_{k+1} < 0\} \cap \{y_k \geq c_{obs}, y_{k+1} \geq c_{obs}\} = D_{c_{obs}}^1 \cup D_{c_{obs}}^2,$$

avec les définitions suivantes pour les événements D_c^1, D_c^2 :

$$\begin{aligned} D_c^1 &= \{x_k < 0, x_{k+1} > 0\} \cap \{y_k \geq c, y_{k+1} \geq c\}, \\ D_c^2 &= \{x_k > 0, x_{k+1} < 0\} \cap \{y_k \geq c, y_{k+1} \geq c\}, \end{aligned}$$

pour tout $c > 0$.

On considère séparément les événements $D_{c_1}^1$ et $D_{c_2}^2$.

• Sur $D_{c_1}^1$, on a les récurrences suivantes:

$$\begin{aligned} x_{k+1} &= (1 + \varepsilon B_-)x_k + \sigma_- \sqrt{\varepsilon} u_k, \\ y_k &= H_- x_k + \sqrt{\varepsilon} v_k, \quad y_{k+1} = H_+ x_{k+1} + \sqrt{\varepsilon} v_{k+1}. \end{aligned}$$

On déduit l'égalité:

$$\frac{1}{\sqrt{\varepsilon}} (H_- y_{k+1} - H_+ (1 + \varepsilon B_-) y_k) = H_+ H_- \sigma_- u_k + H_- v_{k+1} - H_+ (1 + \varepsilon B_-) v_k.$$

Or, d'après (H3') on a $h(x_k) \geq 0$ et, pour ε suffisamment petit, on obtient la majoration

$$\frac{1}{\sqrt{\varepsilon}} |H_- - H_+ (1 + \varepsilon B_-)| c_1 \leq |Z_k|,$$

où

$$Z_k = H_+ H_- \sigma_- u_k + H_- v_{k+1} - H_+ (1 + \varepsilon B_-) v_k$$

suit une loi gaussienne $\mathcal{N}(0, H_+^2 H_-^2 \sigma_-^2 + H_-^2 + H_+^2 (1 + \varepsilon B_-)^2)$.

Soit $\{Z_k^0\}$ le processus normalisé associé à $\{Z_k\}$. Pour un risque α donné ⁽²⁾, on détermine $\lambda > 0$ tel que

$$P(|Z_k^0| \geq \lambda) \leq \alpha,$$

²Généralement on prend $\alpha = 0.05$

par inversion de la fonction de répartition de la loi normale centrée réduite. On obtient ainsi l'expression de la constante c_1 :

$$c_1 = \lambda \sqrt{\varepsilon} \frac{\sqrt{H_+^2 H_-^2 \sigma_-^2 + H_-^2 + H_+^2 (1 + \varepsilon B_-)^2}}{|H_- - H_+ (1 + \varepsilon B_-)|}.$$

• Sur $D_{c_2}^2$, un raisonnement similaire au précédent nous amène à considérer une constante c_2 donnée par:

$$c_2 = \lambda \sqrt{\varepsilon} \frac{\sqrt{H_+^2 H_-^2 \sigma_+^2 + H_+^2 + H_-^2 (1 + \varepsilon B_+)^2}}{|H_+ - H_- (1 + \varepsilon B_+)|}.$$

Dans les applications numériques, pour une probabilité d'erreur α donnée, nous prendrons $c_{obs} = \max(c_1, c_2)$.

Remarque 3.2 Dans les calculs de la constante c_{obs} , la prise en compte de la dérive introduit uniquement des termes d'ordre $\mathcal{O}(\varepsilon^{3/2})$ que nous négligerons devant les termes d'ordre $\mathcal{O}(\sqrt{\varepsilon})$.

3.2 Test basé sur la sortie d'un filtre de Kalman-Bucy

On se propose de décrire un test capable de détecter les changements de signe éventuels du processus $\{x_k\}$ à l'aide de la sortie obtenue par un des deux filtres de Kalman-Bucy, par exemple $\{\hat{x}_k^+\}$. La construction d'un tel test se justifie par le fait que " $h(x) = 0$ ssi $x = 0$ " et par la proposition suivante:

Proposition 3.3 Soit $0 < a < b$. Pour tout $\theta > 0$, il existe $\beta > 0$, $\varepsilon_0 > 0$ tels que, pour $\varepsilon \in (0, \varepsilon_0]$, si on définit $i_0 = \lfloor a/\varepsilon \rfloor$ et $m = \lfloor (b-a)/\varepsilon \rfloor$, on a:

$$P(\{\max_{k=i_0, \dots, i_0+m} |h(x_k) - H_+ \hat{x}_k^+| > \theta\}) \leq e^{-\beta/\sqrt{\varepsilon}}.$$

En vue de la démonstration de cette proposition on établit d'abord le lemme suivant:

Lemme 3.4 Pour tout $\tau > 0$, il existe $\varepsilon_0 > 0$ et $c, C > 0$ tels que, pour tout $\varepsilon \in (0, \varepsilon_0]$, $m = \lfloor \tau/\varepsilon \rfloor$,

$$\mathbf{E}(\exp \max_{k=0, \dots, m} c x_k^2) \leq C.$$

Preuve

La preuve de ce lemme sera présentée en 3 étapes (cf. [Roubaud] pour la preuve analogue en temps continu).

Etape 1:

On démontre le résultat suivant:

Lemme 3.5 On définit le processus $\{z_k\}$ par

$$z_{k+1} = (1 + \varepsilon B)z_k + \sqrt{\varepsilon} \sigma u_k, \quad (9)$$

où z_0 est une variable aléatoire indépendante du processus $\{u_k\}$ et il existe des constantes $c_0, C_0 > 0$ telles que

$$\mathbb{E}(\exp c_0 z_0^2) < C_0. \quad (10)$$

Alors, pour tout $\tau > 0$, il existe ε_0 et $c, C > 0$ tels que, pour tout $\varepsilon \in (0, \varepsilon_0]$, $m = \lceil \tau/\varepsilon \rceil$,

$$\mathbb{E}(\exp \max_{k=0, \dots, m} c z_k^2) \leq C.$$

Preuve

Nous réécrivons l'expression de z_{k+1} en mettant en évidence une partie dépendante de z_0 et une partie martingale:

$$z_{k+1} = (1 + \varepsilon B)^{k+1} z_0 + (1 + \varepsilon B)^k M_k,$$

avec

$$M_k = \sigma \sqrt{\varepsilon} \sum_{j=0}^k (1 + \varepsilon B)^{-j} u_j.$$

Puisque z_0 et $\{M_k\}$ sont indépendants nous avons:

$$z_{k+1}^2 = (1 + \varepsilon B)^{2(k+1)} z_0^2 + (1 + \varepsilon B)^{2k} M_k^2 \leq e^{2\tau B} (z_0^2 + M_k^2)$$

D'après (10), il est suffisant de montrer qu'il existe $\varepsilon_0, c, C > 0$ tels que

$$\forall \varepsilon < \varepsilon_0, \mathbb{E}(\exp \max_{k=0, \dots, m-1} c M_k^2) \leq C.$$

Or $\{M_k\}$ est un processus gaussien centré de variance

$$\mathbb{E}[M_k^2] = \sigma^2 \frac{(1 + \varepsilon B) - (1 + \varepsilon B)^{-2k}}{B(2 + \varepsilon B)}$$

donc

$$\text{si } B = 0, \mathbb{E}[M_k^2] \leq k \varepsilon \sigma^2$$

$$\text{si } B > 0, \mathbf{E}[M_k^2] \leq \sigma^2 \frac{(1 + \varepsilon B)^2}{B(2 + \varepsilon B)}$$

$$\text{si } B < 0, \mathbf{E}[M_k^2] \leq \sigma^2 \frac{-(1 + \varepsilon B)^2 + \exp(2k\varepsilon B + k\varepsilon^2 B^2)}{-B(2 + \varepsilon B)}.$$

Dans l'intervalle fini $[0, \tau]$ il existe donc ε_0 et une constante $C > 0$ tels que

$$\forall \varepsilon < \varepsilon_0, \max_{k=0, \dots, m} \mathbf{E}[M_k^2] \leq C.$$

d'où on déduit ε_0 et des constantes $c, C > 0$ tels que

$$\forall \varepsilon < \varepsilon_0, \mathbf{E}(\exp c M_{m-1}^2) \leq C.$$

Vu que $\{\exp c M_k^2\}$ est une sous-martingale on peut appliquer la propriété

$$\mathbf{E}(\max_{k=0, \dots, m-1} \exp c M_k^2) \leq \mathbf{E}(\exp 2c M_{m-1}^2).$$

Alors, il existe ε_0 et des constantes $c, C > 0$ tels que

$$\forall \varepsilon < \varepsilon_0, \mathbf{E}(\max_{k=0, \dots, m-1} \exp c M_k^2) \leq C,$$

d'où le lemme. ■

Etape 2:

On introduit les notations:

$$B = \sup(|B_+|, |B_-|)$$

$$\sigma = \sup(|\sigma_+|, |\sigma_-|)$$

On démontre que les moments d'ordre $2p, p \geq 1$, du processus $\{x_k\}$ peuvent être majorés par les moments d'ordre $2p$ du processus $\{z_k\}$ défini dans le lemme précédent.

En effet, d'après l'équation d'état, on a

$$x_{k+1}^{2p} = \sum_{j=0}^{2p} C_j^{2p} [x_k + \varepsilon b(x_k)]^j [\sqrt{\varepsilon} \sigma(x_k) u_k]^{2p-j}.$$

Le fait que $\{u_k\}$ est un bruit blanc et que u_k est indépendant de x_k entraîne l'égalité

$$\mathbf{E}[x_{k+1}^{2p}] = \sum_{j=0; j \text{ pair}}^{2p} C_j^{2p} \mathbf{E}[(x_k + \varepsilon b(x_k))^j (\sqrt{\varepsilon} \sigma(x_k))^{2p-j}] \mathbf{E}[u_k^{2p-j}].$$

On obtient la majoration

$$\mathbf{E}[x_{k+1}^{2p}] \leq \sum_{j=0; j \text{ pair}}^{2p} C_j^{2p} (\sqrt{\varepsilon} \sigma)^{2p-j} \frac{(2p-j)!}{2^{p-j/2} (p-j/2)!} (1 + \varepsilon B)^j \mathbf{E}[x_k^j].$$

Or, de la définition du processus $\{z_k\}$, vient que

$$\mathbf{E}[z_{k+1}^{2p}] = \sum_{j=0; j \text{ pair}}^{2p} C_j^{2p} (\sqrt{\varepsilon} \sigma)^{2p-j} \frac{(2p-j)!}{2^{p-j/2} (p-j/2)!} (1 + \varepsilon B)^j \mathbf{E}[z_k^j].$$

On prend $z_0 = |x_0|$. Alors

$$\mathbf{E}[x_k^{2p}] \leq \mathbf{E}[z_k^{2p}], \quad k = 0, 1, \dots, [T/\varepsilon].$$

Le lemme 3.5 entraîne l'existence de ε_0 , c , $C > 0$ tels que

$$\forall \varepsilon < \varepsilon_0, \quad \max_{k=0, \dots, m} \mathbf{E}(\exp c x_k^2) \leq C. \quad (11)$$

Etape 3:

Nous utilisons la majoration (11) pour obtenir le résultat prétendu, le lemme 3.4.

D'après la formule de Taylor, on a:

$$\begin{aligned} \exp c x_{j+1}^2 &= \exp c x_j^2 + 2 c x_j \exp c x_j^2 (x_{j+1} - x_j) \\ &\quad + 2 c \exp c \theta_j^2 (1 + 2 c \theta_j^2) (x_{j+1} - x_j)^2. \end{aligned}$$

On obtient l'expression

$$\begin{aligned} \exp c x_{k+1}^2 &= \exp c x_0^2 + 2 c \sum_{j=0}^k x_{k-j} \exp(c x_{k-j}^2) (x_{k+1-j} - x_{k-j}) \\ &\quad + 2 c \sum_{j=0}^k x_{k-j} \exp(c \theta_{k-j}^2) (1 + 2 c \theta_{k-j}^2) (x_{k+1-j} - x_{k-j})^2. \end{aligned}$$

Puisque $|\theta_k| \leq |x_k| + |x_{k+1}|$, on a la majoration

$$\begin{aligned} &\mathbf{E}[\max_{k=0, \dots, m-1} \exp c x_{k+1}^2] \\ &\leq \mathbf{E}[\exp c x_0^2] + 2 c \varepsilon B \sum_{j=0}^{m-1} \mathbf{E}[x_{k-j}^2 \exp(c x_{k-j}^2)] \end{aligned}$$

$$\begin{aligned}
& + 2c \mathbf{E} \left[\max_{k=0, \dots, m-1} \sqrt{\varepsilon} \left(\sum_{j=0}^k \exp(c \theta_{k-j}^2) x_{k-j} \sigma(x_{k-j}) u_{k-j} \right) \right] \\
& + 4c\varepsilon \sum_{j=0}^{m-1} \mathbf{E} \left[\exp(2c(x_{k-j}^2 + x_{k+1-j}^2)) (1 + 4c(x_{k+1-j}^2 + x_{k-j}^2)) (\varepsilon B^2 x_{k-j}^2 + \sigma^2 u_{k-j}^2) \right].
\end{aligned} \tag{12}$$

Le processus

$$M_k = \sqrt{\varepsilon} \sum_{j=0}^k \exp(c x_{k-j}^2) x_{k-j} \sigma(x_{k-j}) u_{k-j}$$

est une \mathcal{F}_k -martingale, où $\mathcal{F}_k = \sigma(x_0, u_0, \dots, u_k)$.

On utilise une propriété des martingales:

$$\mathbf{E} \left[\max_{k=0, \dots, m-1} M_k^2 \right] \leq 4 \mathbf{E}[M_{m-1}^2].$$

Or u_k et x_k sont indépendants et x_k est \mathcal{F}_{k-1} -mesurable donc on a

$$\mathbf{E}[M_{m-1}^2] \leq m \varepsilon \sigma \max_{j=0, \dots, m-1} \mathbf{E}[\exp(4c x_{m-1-j}^2)]^{1/2} \mathbf{E}[x_{m-1-j}^4]^{1/2} \leq c_\tau,$$

d'après (11).

Il existe alors $\varepsilon_0 > 0$ et $c, C > 0$ tels que, pour tout $\varepsilon \in (0, \varepsilon_0]$,

$$\mathbf{E} \left[\max_{k=0, \dots, m-1} \sqrt{\varepsilon} \left(\sum_{j=0}^k \exp(c x_{k-j}^2) x_{k-j} \sigma(x_{k-j}) u_{k-j} \right) \right] \leq C.$$

On utilise l'inégalité de Cauchy-Schwarz, le résultat (11) et l'hypothèse (H2) dans l'inégalité (12) pour déduire le lemme. ■

Preuve (de la proposition 3.3)

Nous utilisons la notion de sous-différentiel de h .

Puisque la fonction h est convexe, il existe ρ_k appartenant au sous-différentiel de h à un point θ_k tel que

$$h(x_{k+1}) = h(x_k) + \rho_k(x_{k+1} - x_k)$$

et

$$|\rho_k| \leq \sup(|H_+|, |H_-|) \triangleq H.$$

On définit

$$Z_k \triangleq h(x_k) - H_+ \hat{x}_k^+.$$

D'après l'équation du filtre \hat{x}_k^+ et l'équation d'état, on obtient l'expression ré-
cursive:

$$Z_{k+1} = \delta Z_k + \varepsilon b(x_k) \rho_k (1 - H_+ K_+) + \sqrt{\varepsilon} \sigma(x_k) \rho_k (1 - H_+ K_+) u_k \\ + \sqrt{\varepsilon} H_+ K_+ v_{k+1},$$

où

$$\delta \triangleq (1 + \varepsilon B_+)(1 - H_+ K_+).$$

On remarque que $0 < \delta < 1$. On utilise cette récurrence jusqu'au terme Z_{i_0} . On établit ainsi la majoration:

$$|Z_{k+1}| \leq \exp\{-c(k+1)\} |Z_{i_0}| + C \varepsilon \sum_{j=0}^{k-i_0} \delta^j |x_{k-j}| \\ + C \sqrt{\varepsilon} \sum_{j=0}^{k-i_0} \delta^j |u_{k-j}| + C \sqrt{\varepsilon} \sum_{j=0}^{k-i_0} \delta^j |v_{k+1-j}|. \quad (13)$$

On traite les quatre termes dans cette expression séparément.

D'après l'inégalité de Bienaymé-Tchebichev et le fait que $\mathbb{E}[|Z_{i_0}|] < C$, on obtient

$$P\left(|Z_{i_0}| > \frac{\theta}{4} e^{c i_0}\right) \leq \frac{1}{\theta} e^{-c a/\varepsilon}.$$

D'autre part, en utilisant le lemme 3.4, on obtient la majoration

$$P\left(\varepsilon \sum_{j=0}^{k-i_0} \delta^j |x_{k-j}| > \frac{\theta}{4C}\right) \leq e^{-\theta(1-\delta)/(4C\varepsilon)}$$

Le lemme 2.3. dans [Roubaud] entraîne la majoration

$$P\left(\max_{k=i_0, \dots, i_0+m-1} |u_k| > \frac{\theta(1-\delta)}{4C\sqrt{\varepsilon}}\right) \leq \frac{h-a}{\varepsilon} e^{-c\theta/\varepsilon},$$

donc

$$P\left(\sqrt{\varepsilon} \sum_{j=0}^{k-i_0} \delta^j |u_{k-j}| > \frac{\theta}{4C}\right) \leq e^{-c/\varepsilon},$$

pour ε "suffisamment petit".

De même pour le 4^{me} terme à droite dans l'expression (13):

$$P\left(\sqrt{\varepsilon} \sum_{j=0}^{k-i_0} \delta^j |v_{k-j}| > \frac{\theta}{4C}\right) \leq e^{-c/\varepsilon}.$$

De l'expression (13), en utilisant les 4 majorations ci-dessus, on obtient le résultat. ■

Soit c_{FK} une constante strictement positive à déterminer.

On définit l'évènement test suivant:

$$\hat{C} = \{|\hat{x}_k^+| \geq c_{FK}; k = i_0, i_0 + 1, \dots, i_0 + m\}.$$

Comme conséquence de la proposition 3.3, on obtient le théorème suivant:

Théorème 3.6 *Il existe $\beta > 0$, $\varepsilon_0 > 0$ tels que, pour $\varepsilon \in (0, \varepsilon_0]$, on a:*

$$P((A_+ \cup A_-)^c | \hat{C}) \leq e^{-\beta/\sqrt{\varepsilon}}.$$

Preuve La preuve de ce théorème suit pas à pas celle du résultat équivalent en temps continu (cf. [Roubaud] [Théorème 3.2.]). ■

Détermination de la constante c_{FK} :

On cherche une constante c_{FK} telle que la probabilité

$$P(\{x_k x_{k+1} \leq 0\} | \{|\hat{x}_k^+| \geq c_{FK}, |\hat{x}_{k+1}^+| \geq c_{FK}\})$$

soit faible.

On considère y_k comme une approximation de $h(x_k)$. Dans l'étude précédente, pour une probabilité d'erreur α donnée, on a déterminé une constante c_{obs} permettant de tester au vu de y_k l'occurrence d'un changement du signe de x_k . Or, le changement de signe de x_k est directement lié au fait que $h(x_k)$ devient "petit". Si $H_+ \hat{x}_k^+$ était une meilleure approximation de $h(x_k)$ que y_k , tout au moins aux alentours de zéro, la probabilité d'erreur du test " $\hat{x}_k^+ \geq c_{obs}/H_+$ " serait minorée par α . Cependant, nous ne savons pas si cette condition est vérifiée mais, sur l'évènement A_+ , le filtre \hat{x}_m^+ est une meilleure approximation de x_m que y_m/H_+ . En tout cas, la proposition 3.3 nous assure que l'écart $|h(x_k) - H_+ \hat{x}_k^+|$ est petit avec une probabilité proche de 1. Il semble alors raisonnable de penser que le test " $\hat{x}_k^+ \geq c_{obs}/H_+$ " permettra de détecter convenablement les passages de x_k par 0. Nous proposons donc pour constante

$$c_{FK} = c_{obs}/|H_+|.$$

Le même type de raisonnement est valable quand le test est appliqué sur le filtre \hat{x}_k^- au lieu du filtre \hat{x}_k^+ . Nous prenons alors

$$c_{FK} = c_{obs} / |H_-| .$$

Sans perte de généralité on considèrera par la suite un seul intervalle $[a, b]$ représentant l'intervalle de temps durant lequel il n'y a pas de passage à 0, avec une probabilité d'erreur donnée, i_0 et m étant définis comme en début de cette section.

4 Sous l'hypothèse (HD1): décider du signe de x_k

Dans cette section, nous présentons deux tests permettant de choisir entre les événements A_+ et A_- , dans le cas où le problème de filtrage linéaire par morceaux vérifie l'hypothèse (HD1). Ces tests ont été étudiés dans [FJSZ]. Le problème analogue en temps continu a été traité dans [FJP]. Le premier test considéré, dit de la variation quadratique, n'est applicable que sous l'hypothèse (HD1). Le second test, du type rapport de vraisemblance sur les sorties du filtre de Kalman-Bucy, pourra être adapté au cas où l'hypothèse (HD2) est vérifiée (cf. paragraphe 5.2). Le premier test sera donc introduit de façon brève. Le deuxième test sera décrit de façon plus détaillée, puisque la même idée sera utilisée par la suite.

Dans [FJSZ], les auteurs ont mis en œuvre ces tests sous deux formes: test à taille d'échantillon fixée et test séquentiel. D'après leur étude, les tests séquentiels semblent plus intéressants que ceux à taille d'échantillon fixée, vis à vis du critère du "temps moyen pour prendre une décision". Nous ne présenterons donc que les tests séquentiels.

4.1 Test de la variation quadratique

On considère un test d'hypothèses basé sur la variation quadratique de la suite d'observations $\{y_k\}$ pour décider entre les deux alternatives " A_+ " et " A_- ". Cette décision est prise sur N observations ($0 < N \leq m$) dans un intervalle de monotonie obtenu par un des tests introduits dans la section 3. On considère N comme étant un temps d'arrêt.

Sans restreindre la généralité du problème, on suppose que $B_+ = B_- = B$.

Remarque 4.1 Le fait de considérer B_+ et B_- distincts a pour unique conséquence l'introduction de termes négligeables dans les expressions, la procédure restant essentiellement la même.

Notons:

$$\begin{aligned}\Delta_k &= y_{k+1} - (1 + B\varepsilon) y_k, \\ \Upsilon_+^2 &= H_+^2 \sigma_+^2 + 1 + (1 + B\varepsilon)^2, \\ \Upsilon_-^2 &= H_-^2 \sigma_-^2 + 1 + (1 + B\varepsilon)^2.\end{aligned}$$

Alors

- si $x_k > 0$ et $x_{k+1} < 0$, la variable aléatoire Δ_k suit une loi gaussienne de moyenne nulle et variance $\Upsilon_+^2 \varepsilon$

- si $x_k < 0$ et $x_{k+1} < 0$, la variable aléatoire Δ_k suit une loi gaussienne de moyenne nulle et variance $\Upsilon_-^2 \varepsilon$.

L'hypothèse de détectabilité (HD1) implique

$$\Upsilon_+^2 \neq \Upsilon_-^2.$$

Le problème de décider entre les alternatives " A_+ " et " A_- " conduit à un test d'hypothèses sur ces variances.

On définit le rapport logarithmique

$$Z_k = \ln \frac{\mathcal{N}(\Delta_k, 0, \Upsilon_+^2 \varepsilon)}{\mathcal{N}(\Delta_k, 0, \Upsilon_-^2 \varepsilon)},$$

où $\mathcal{N}(x, \mu, \gamma^2)$ représente la densité de la loi gaussienne de paramètres μ et γ^2 et la série

$$S_n = \sum_{k=i_0}^{i_0+n} Z_k,$$

où i_0 et m sont définis dans la section 3.

On obtient les expressions suivantes:

$$Z_k = \log \frac{\Upsilon_-}{\Upsilon_+} + \frac{1}{2} \left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2} \right) \frac{\Delta_k^2}{\varepsilon}$$

et

$$S_n = (n+1) \log \frac{\Upsilon_-}{\Upsilon_+} + \frac{1}{2} \left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2} \right) \sum_{k=i_0}^{i_0+n} \frac{\Delta_k^2}{\varepsilon}.$$

Soient $l_1 > 0$ et $l_2 > 0$ fixés. On introduit le temps d'arrêt N^* :

$$N^* = \inf\{n > 0 : S_n < -l_1 \text{ ou } S_n > l_2\} \wedge m.$$

On décrit le test comme suit: si $S_{N^*} \geq l_2$ on accepte " A_+ " et si $S_{N^*} \leq -l_1$ on accepte " A_- "; sinon on dira que le test ne permet pas de décider.

Il reste à proposer une méthode de détermination des bornes l_1 et l_2 . On approche le processus discret $\{\varepsilon S_n\}$ par un certain processus de diffusion $\{\zeta_t\}$ dont on écrit les équations sous " A_+ " et sous " A_- ". On obtient ainsi, par la formule de Dynkin, des équations différentielles ordinaires permettant de calculer les temps moyens pour prendre une décision $E(T_+^*)$ et $E(T_-^*)$ et pour les probabilités d'erreur p_+ et p_- (cf. [Fleming-Rishel]).

On introduit les notations

$$\begin{aligned}\mu_+ &= \log \frac{\Upsilon_-}{\Upsilon_+} + \frac{1}{2} \left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2} \right) \Upsilon_+^2 \\ \mu_- &= \log \frac{\Upsilon_-}{\Upsilon_+} + \frac{1}{2} \left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2} \right) \Upsilon_-^2 \\ \gamma_+ &= \frac{1}{2} \left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2} \right) \sqrt{\Upsilon_+^4 + 2(1+B\varepsilon)^2} \\ \gamma_- &= \frac{1}{2} \left(\frac{1}{\Upsilon_-^2} - \frac{1}{\Upsilon_+^2} \right) \sqrt{\Upsilon_-^4 + 2(1+B\varepsilon)^2}.\end{aligned}$$

La résolution des équations nous donne comme solution:

$$p_+ = \frac{1 - e^{-\mu_+/\gamma_+^2 l_2}}{e^{\mu_+/\gamma_+^2 l_1} - e^{-\mu_+/\gamma_+^2 l_2}} \quad (14)$$

$$p_- = \frac{1 - e^{-\mu_-/\gamma_-^2 l_1}}{e^{-\mu_-/\gamma_-^2 l_2} - e^{\mu_-/\gamma_-^2 l_1}} \quad (15)$$

$$E(T_+^*) = \left(\frac{l_2}{\mu_+} - p_+ \frac{l_2 + l_1}{\mu_+} \right) \varepsilon$$

$$E(T_-^*) = \left(-\frac{l_1}{\mu_-} + p_- \frac{l_2 + l_1}{\mu_-} \right) \varepsilon.$$

où

- p_+ est "la probabilité de refuser A_+ à tort".
- p_- est "la probabilité de refuser A_- à tort".
- T_+^* et T_-^* sont les temps d'arrêt

$$T_+^* = \inf\{t : \zeta_t \geq \varepsilon l_2 \text{ ou } \zeta_t \leq -\varepsilon l_1\} \text{ , dans le cas positif.}$$

et

$$T_-^* = \inf\{t : \zeta_t \geq \varepsilon l_2 \text{ ou } \zeta_t \leq -\varepsilon l_1\} \text{ , dans le cas négatif .}$$

Pour des probabilités d'erreur p_+ et p_- données ⁽³⁾, les expressions (14) et (15) permettent de calculer les bornes l_1 et l_2 .

³Généralement on prend $p_+ = p_- = 0.05$.

4.2 Test du rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy

On décrit un autre test du type rapport de vraisemblance pour décider entre les deux alternatives " A_+ " et " A_- ". L'idée de ce test est la suivante:

Soit i_1 un entier, $i_0 < i_1 < i_0 + m$. Intuitivement, on s'attend à ce que $\hat{h}_k = E(h(X_k)|\mathcal{Y}_0^k)$ soit très proche de $H_+ \hat{x}_k^+$ sur A_+ et de $H_- \hat{x}_k^-$ sur A_- , pour $i_1 \leq k \leq i_0 + m$. On a

$$\begin{aligned} \text{sur } A_+ : \quad y_k &= H_+(1 + \varepsilon B_+) \hat{x}_{k-1}^+ + \nu_k^+, \\ \text{sur } A_- : \quad y_k &= H_-(1 + \varepsilon B_-) \hat{x}_{k-1}^- + \nu_k^-, \end{aligned}$$

où $\{\nu_k^+\}$ et $\{\nu_k^-\}$ peuvent être considérés comme des bruits blancs gaussiens dits "processus d'innovation" (cf. section 2).

On pose

$$Z_k = H_+(1 + \varepsilon B_+) \hat{x}_k^+ - H_-(1 + \varepsilon B_-) \hat{x}_k^-$$

et on prend pour statistique de test le rapport logarithmique suivant:

$$\hat{L}_n = \frac{1}{\varepsilon} \left[\sum_{k=i_1}^{i_1+n} Z_k y_{k+1} - \frac{1}{2} \sum_{k=i_1}^{i_1+n} \left((H_+(1 + \varepsilon B_+) \hat{x}_k^+)^2 - (H_-(1 + \varepsilon B_-) \hat{x}_k^-)^2 \right) \right],$$

avec $i_1 \leq n \leq m$.

La statistique \hat{L}_n se réécrit sous la forme suivante:

$$\begin{aligned} \text{Sur } A_+ : \hat{L}_n &= +\frac{1}{2\varepsilon} \sum_{k=i_1}^{i_1+n} Z_k^2 + \frac{1}{\varepsilon} \sum_{k=i_1}^{i_1+n} Z_k \nu_{k+1}^+, \\ \text{Sur } A_- : \hat{L}_n &= -\frac{1}{2\varepsilon} \sum_{k=i_1}^{i_1+n} Z_k^2 + \frac{1}{\varepsilon} \sum_{k=i_1}^{i_1+n} Z_k \nu_{k+1}^-. \end{aligned}$$

On note \hat{R}_n le terme de précision du test:

$$\hat{R}_n = \frac{1}{\varepsilon} \sum_{k=i_1}^{i_1+n} Z_k^2.$$

Pour R_n "suffisamment grand", le signe de \hat{L}_n permet de choisir entre les alternatives " A_- " et " A_+ ".

Soient $l_1 > 0$ et $l_2 > 0$ des bornes fixées. On introduit le temps d'arrêt N^* :

$$N^* = \inf\{n \geq i_1 : \hat{L}_n \geq l_2 \text{ ou } \hat{L}_n \leq -l_1\} \wedge m.$$

On applique la règle de décision qui suit.

Règle de décision:

- Si $\hat{L}_{N^*} \geq l_2$, on décide " A_+ ".
- Si $\hat{L}_{N^*} \leq -l_1$, on décide " A_- ".
- Sinon, on ne peut pas décider.

Il nous reste à calculer les bornes l_1 et l_2 pour une probabilité d'erreur donnée α et à estimer les temps moyens pour atteindre une décision. Le raisonnement présenté par la suite constitue une justification purement heuristique, sans la prétention d'être une preuve rigoureuse.

Si on se place dans le cadre de l'hypothèse " A_+ ", d'après les équations des filtres de Kalman-Bucy (7) et (8), on a:

$$\begin{aligned} Z_{k+1} = & (1 + \varepsilon B_-)(1 - H_- K_-) Z_k + \varepsilon (1 + \varepsilon B_+)(B_+ - B_-) H_+ \hat{x}_k^+ \\ & + (H_+ K_+ (1 + \varepsilon B_+) - H_- K_- (1 + \varepsilon B_-)) \nu_{k+1}^+. \end{aligned} \quad (16)$$

Vu que les filtres en question sont à mémoire courte, d'après cette expression, l'hypothèse d'indépendance sur A^+ de ν_k^+ et ν_{k+1}^+ , pour i_0, \dots, m entraîne asymptotiquement "l'indépendance" de Z_k et ν_{k+1}^+ , à partir d'un certain instant $i_1 > i_0$.

Remarque 4.2 Pour ε "assez petit", on a la majoration

$$(1 + \varepsilon B_-)(1 - H_- K_-) < 1.$$

On obtient alors que

- si $B_+ = B_- \triangleq B$, dans une situation asymptotique,

$$E[Z^2; A_+] = \varepsilon \frac{(1 + \varepsilon B)^2 (H_+ K_+ - H_- K_-)^2}{1 - (1 + \varepsilon B)^2 (1 - H_- K_-)^2} \vartheta_+.$$

- si $B_+ < 0$, alors $E[\hat{x}_k^{+2}; A_+]$ admet une solution d'équilibre:

$$E[\hat{x}_k^{+2}; A_+] = \frac{K_+^2}{|B_+|(2 + \varepsilon B_+)} \vartheta^+.$$

On utilise l'inégalité

$$\mathbf{E}[Z_k \hat{x}_k^+; A_+] \leq \frac{1}{2\sqrt{\varepsilon}} \mathbf{E}[Z_k^2; A_+] + \frac{\sqrt{\varepsilon}}{2} \mathbf{E}[\hat{x}_k^+; A_+]$$

pour obtenir l'estimation

$$\begin{aligned} & \mathbf{E}[Z_{k+1}^2; A_+] \\ & \leq [(1 + \varepsilon B_-)^2 (1 - H_- K_-)^2 + \sqrt{\varepsilon} H_+ |B_+ - B_-| (1 - H_- K_-) (1 + \varepsilon B_+) (1 + \varepsilon B_-)] \mathbf{E}[Z_k^2; A_+] \\ & \quad + \varepsilon^{3/2} (1 + \varepsilon B_+) (1 + \varepsilon B_-) |B_+ - B_-| (1 - H_- K_-) H_+ \mathbf{E}[\hat{x}^{+2}; A_+] \\ & \quad + \varepsilon [H_+ K_+ (1 + \varepsilon B_+) - H_- K_- (1 + \varepsilon B_-)]^2 \vartheta^+ \\ & \quad + \varepsilon^2 (1 + \varepsilon B_+)^2 (B_+ - B_-)^2 H_+^2 \mathbf{E}[\hat{x}^{+2}; A_+]. \end{aligned}$$

Remarque 4.3 Pour ε "assez petit", on a la majoration

$$\begin{aligned} & (1 + \varepsilon B_-)^2 (1 - H_- K_-)^2 + \sqrt{\varepsilon} H_+ |B_+ - B_-| (1 - H_- K_-) (1 + \varepsilon B_+) (1 + \varepsilon B_-) \\ & < 1 - H_- K_- (2 - H_- K_-) + c \sqrt{\varepsilon} < 1. \end{aligned}$$

Dans une situation stationnaire, on a l'estimation

$$\mathbf{E}[Z^2; A_+] \leq \Lambda_+ \varepsilon + c \varepsilon^{3/2},$$

où

$$\Lambda_+ = \frac{(H_+ K_+ - H_- K_-)^2}{H_- K_- (2 - H_- K_-)} \vartheta^+.$$

De plus, pour un certain n_0 , on peut considérer Z_k comme une combinaison linéaire de $\nu_{k-n_0}^+, \dots, \nu_k^+$. Le processus $\xi_k = Z_k \nu_{k+1}^+$ est alors approximativement un processus mélangeant. En conséquence, on approche $\varepsilon \hat{L}_n$ par le processus de diffusion $\{\zeta_t; t = \varepsilon n\}$ solution de l'équation différentielle stochastique

$$d\zeta_t = \Lambda_+/2 dt + \sqrt{\varepsilon \Lambda_+ \vartheta^+} dW_t^+.$$

Comme pour le test de la variation quadratique, on obtient la probabilité d'erreur $p_+ = P(\{\zeta_{T_+^*} = -\varepsilon l_1\} | A_+)$:

$$p_+ = \frac{1 - e^{-\eta_+ l_2}}{e^{\eta_+ l_1} - e^{-\eta_+ l_2}},$$

où T_+^* est le temps d'arrêt défini par

$$T_+^* = \inf\{t : \zeta_t \geq \varepsilon l_2 \text{ ou } \zeta_t \leq -\varepsilon l_1\}$$

et $\eta_+ = 1/\vartheta^+$.

Si on se place dans le cadre de l'hypothèse " A_- ", par un raisonnement identique au précédent, on obtient la probabilité d'erreur $p_- = P(\{\zeta_{T_-} = \varepsilon l_2\} | A_-)$. Elle est donnée par:

$$p_- = \frac{1 - e^{-\eta_- l_1}}{e^{\eta_- l_2} - e^{-\eta_- l_1}},$$

avec $\eta_- = 1/\vartheta^-$.

Pour des valeurs de p_+ et p_- données ⁽⁴⁾, on détermine les bornes de précision l_1 et l_2 et on estime les temps moyens pour prendre une décision:

$$\begin{aligned} E(T_+^*) &= 2\varepsilon \left(\frac{l_2}{\Lambda_+} - p_+ \frac{l_1 + l_2}{\Lambda_+} \right), \\ E(T_-^*) &= 2\varepsilon \left(\frac{l_1}{\Lambda_-} - p_- \frac{l_1 + l_2}{\Lambda_-} \right). \end{aligned}$$

Remarque 4.4 Dans [FJSZ], les auteurs constatent qu'en pratique le temps moyen pour prendre une décision est plus long pour le test de rapport de vraisemblance que pour le test de variation quadratique.

⁴En général on prend $p_- = p_+ = 0.05$.

5 Sous l'hypothèse (HD2): décider du signe de

x_k

Sous l'hypothèse (HD2): $\Gamma^2 = (H_- \sigma_-)^2 = (H_+ \sigma_+)^2$ et $B_- \neq B_+$, le test de la variation quadratique n'est pas applicable. En remplacement nous considérons un test du type rapport de vraisemblance sur les observations $\{y_k\}$. Puis nous adapterons, sous cette hypothèse, le test du rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy, présenté dans le paragraphe précédent sous l'hypothèse (HD1).

Les calculs qui suivent utilisent le fait qu'une situation stationnaire est nécessairement atteinte, raison pour laquelle nous avons formulé l'hypothèse supplémentaire (H4), décrite dans l'introduction. Nous l'utiliserons par la suite.

5.1 Test du rapport de vraisemblance sur les accroissements des observations

Ce test est déduit de l'étude du cas $\varepsilon = 0$.

5.1.1 Le cas $\varepsilon = 0$

On considère le système suivant:

$$\begin{cases} x_{k+1} = x_k + \Delta t b(x_k) + \sqrt{\Delta t} \sigma(x_k) u_k, \\ \tilde{y}_k = h(x_k). \end{cases} \quad (17)$$

L'équation d'état de ce système généralise celle du système discret (4) à un pas de temps Δt .

Le problème est de déterminer le signe de x_k pour $k = i_0, i_0 + 1, \dots, i_0 + m$. Le test d'hypothèses est fait sur les observations $\{\tilde{y}_k, i_0 + 1 \leq k \leq i_0 + m\}$.

On considère la filtration $(\mathcal{F}_{i_0}^k)_k$, où $\mathcal{F}_{i_0}^k = \sigma\{x_{i_0}, u_{i_0}, \dots, u_k\}$, $i_0 \leq k \leq i_0 + m$.

$$\text{Sur } A_+ : \quad \tilde{y}_{k+1} = (1 + \Delta t B_+) \tilde{y}_k + \sqrt{\Delta t} \Gamma u_k^+. \quad (18)$$

où $\{u_k^+; k \geq i_0\}$ est un $\mathcal{F}_{i_0}^k$ -bruit blanc gaussien pour une probabilité P^+ telle que

$$P^+(D) = P(D), \forall D \subset A_+.$$

$$\text{Sur } A_- : \quad \tilde{y}_{k+1} = (1 + \Delta t B_-) \tilde{y}_k + \sqrt{\Delta t} \Gamma u_k^-. \quad (19)$$

où $\{u_k^-; k \geq i_0\}$ est un $\mathcal{F}_{i_0}^k$ -bruit blanc gaussien pour une probabilité P^- telle que

$$P^-(D) = P(D), \forall D \subset A_-.$$

Le processus $\{\tilde{y}_k; k \geq i_0+1\}$ est $\mathcal{F}_{i_0}^{k-1}$ -mesurable (i.e. prévisible) et $\sum_{k=i_0}^{i_0+m} \tilde{y}_k^2 < +\infty$ P -p.s.

On utilise une version discrète du théorème de Girsanov (cf. [Milheiro]) et, vu que $B_- \neq B_+$, on prend pour statistique le logarithme du rapport de vraisemblance:

$$L_n = \ln \frac{dP^+}{dP^-} \Big|_{\mathcal{F}_{i_0}^{i_0+n}},$$

où $1 \leq n \leq m-1$. On a la formule:

$$L_n = \frac{1}{\Gamma^2} \left[(B_+ - B_-) \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k (\tilde{y}_{k+1} - \tilde{y}_k) - \frac{1}{2} \Delta t (B_+^2 - B_-^2) \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k^2 \right]. \quad (20)$$

Sur l'évènement A_+ , la statistique L_n se réécrit sous la forme:

$$L_n = +\frac{1}{2} \Delta t \frac{(B_+ - B_-)^2}{\Gamma^2} \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k^2 + \sqrt{\Delta t} \frac{B_+ - B_-}{\Gamma} \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k u_k^+. \quad (21)$$

Sur l'évènement A_- , on a:

$$L_n = -\frac{1}{2} \Delta t \frac{(B_+ - B_-)^2}{\Gamma^2} \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k^2 + \sqrt{\Delta t} \frac{B_+ - B_-}{\Gamma} \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k u_k^-. \quad (22)$$

On pose

$$R_n = \Delta t \frac{(B_+ - B_-)^2}{\Gamma^2} \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k^2.$$

On appelle R_n le terme de précision. Pour R_n grand, L_n/R_n est proche de $\pm 1/2$ et la probabilité d'erreur est faible. Si $L_n \geq 0$ on décide " A_+ " et si $L_n < 0$ on décide " A_- ".

On établit le résultat suivant:

Théorème 5.1 Soit $r > 0$. $\forall n$ tel que $i_0 < n \leq m-1$, on a:

$$\begin{aligned} P(\{L_n < 0\} \cap \{R_n \geq r\} \cap A_+) &\leq e^{-r/8}, \\ P(\{L_n \geq 0\} \cap \{R_n \geq r\} \cap A_-) &\leq e^{-r/8}. \end{aligned}$$

Preuve La démonstration est identique pour les deux inégalités. Considérons par exemple la première.

On introduit la martingale exponentielle

$$Z_k = \exp \left\{ \lambda \sqrt{\Delta t} \frac{B_+ - B_-}{\Gamma} \sum_{i=i_0+1}^{i_0+k} \tilde{y}_i u_i^+ - \frac{1}{2} \lambda^2 \Delta t \frac{(B_+ - B_-)^2}{\Gamma^2} \sum_{i=i_0+1}^{i_0+k} \tilde{y}_i^2 \right\},$$

avec $\lambda < 0$ et on raisonne comme dans le cas continu (cf. [Roubaud]). ■

On décrit deux procédures de mise en oeuvre de ce test sous forme séquentielle.

Première procédure:

Pour r fixé, on introduit le temps d'arrêt N^* :

$$N^* = \inf\{n \geq 1 : R_n \geq r\} \wedge m.$$

- Si $L_{N^*} \geq 0$ on décide " A_+ ",
- Si $L_{N^*} < 0$ on décide " A_- ";
- Sinon on ne peut pas décider.

Nous devons donc proposer une méthode de détermination de cette constante r . On considère le problème sur A_+ et sur A_- et on choisit $r = \max(r_-, r_+)$.

Sur l'évènement A_+ , on pose

$$M_r^+ = \sqrt{\Delta t} \frac{B_+ - B_-}{\Gamma} \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k u_k^+.$$

La statistique L_n se réécrit sous la forme:

$$L_n = \frac{1}{2} R_n + M_n^+.$$

On introduit le temps d'arrêt N_+^* :

$$N_+^* = \inf\{n \geq 1 : R_n \geq r_+\} \wedge m.$$

On a que $R_{N_+^*}$ est approximativement égal à r_+ .

On cherche à déterminer une constante r_+ telle que, pour une probabilité $1 - \alpha$ donnée, le terme de précision soit significatif dans le calcul de $L_{N_+^*}$. Dans d'autres termes, on cherche r_+ telle que:

$$P(\{-M_{N_+^*}^+ > \frac{1}{2} r_+\} | A_+) \leq \alpha. \quad (23)$$

D'après (18), on réécrit \tilde{y}_k sous la forme suivante:

$$\tilde{y}_{k+1} = (1 + \Delta t B_+)^{n_0} \tilde{y}_{k+1 \dots n_0} + \sqrt{\Delta t} \Gamma \sum_{i=0}^{n_0-1} (1 + \Delta t B_+)^i u_{k-i}^+.$$

L'hypothèse (H4) nous permet de considérer \tilde{y}_k comme une combinaison linéaire des variables aléatoires gaussiennes et indépendantes $u_{k-n_0+1}^+, \dots, u_k^+$. Soit $\xi_k = \tilde{y}_k u_k^+$. Le processus $\{\xi_k, i_0 + n_0 - 1 \leq k \leq i_0 + m\}$ est alors approximativement un processus mélangeant. De plus, la valeur stationnaire de $E[\tilde{y}_k^2; A_+]$ est approchée par

$$\Lambda_+ = \frac{\Gamma^2}{-B_+(2 + \Delta t B_+)}.$$

D'après le théorème fonctionnel de limite centrale [Billingsley][th.20.1],

$$\frac{\sqrt{\Delta t}}{\sqrt{\Lambda_+}} \sum_{k=i_0+1}^{i_0+n} \xi_k \approx W_t,$$

où $t = n\Delta t$ et $\{W_t\}$ est un processus de Wiener standard.

On déduit de ce qui précède que M_n^+ suit approximativement une loi normale

$$\mathcal{N}\left(0, n\Delta t \frac{(B_+ - B_-)^2}{-B_+(2 + \Delta t B_+)}\right).$$

Vu que $R_{N_+^*} \approx r_+$, on obtient la relation suivante entre r_+ et N_+^* :

$$r_+ \approx N_+^* \Delta t \frac{(B_+ - B_-)^2}{-B_+(2 + \Delta t B_+)}.$$

Alors l'inégalité (23) devient:

$$P\left(\left\{\frac{-M_{N_+^*}^+}{\sqrt{r_+}} > \frac{1}{2}\sqrt{r_+}\right\} | A_+\right) \leq \alpha.$$

On choisit

$$r_+ = 4\lambda^2,$$

avec λ obtenu par la table de la loi normale centrée réduite.

Sur l'évènement A_- , on prend

$$r_- = 4\lambda^2$$

et on a la relation suivante entre r_- et N_-^* :

$$r_- \approx N_-^* \Delta t \frac{(B_+ - B_-)^2}{-B_-(2 + \Delta t B_-)}.$$

Exemple 5.2 Pour $B_- = -1$, $B_+ = -0.25$ et $\alpha = 0.05$ on obtient $r = 10.89$. Pour les estimations des temps moyens on obtient: $\Delta t N_+^* = 9.66$, $\Delta t N_-^* = 38.52$.

Deuxième procédure:

Il semble que dans le test précédent le temps d'attente pour prendre une décision soit "trop" long. En effet les bornes r_+ et r_- sont déterminées indépendamment l'une de l'autre. Pour réduire cette durée, nous proposons une autre procédure.

Pour des constantes $l_1 > 0$, $l_2 > 0$ à déterminer, on introduit le temps d'arrêt N^* :

$$N^* = \inf\{n > i_0 \text{ t.q. } L_n \geq l_2 \text{ ou } L_n \leq -l_1\} \wedge m.$$

Règle de décision :

- Si $L_{N^*} \geq l_2$ on décide " A_+ ".
- Si $L_{N^*} \leq -l_1$ on décide " A_- ".
- Sinon, on ne peut pas décider.

On utilise à nouveau l'approximation des processus discrets par des diffusions pour déterminer les constantes l_1 et l_2 .

Sur A_+ , on approche L_n par $\{\zeta_t; t = \Delta t n\}$, où $\{\zeta_t\}$ est le processus de diffusion défini par l'équation différentielle stochastique

$$d\zeta_t = \frac{(B_+ - B_-)^2 \Lambda_+}{2 \Gamma^2} dt + \frac{\sqrt{\Lambda_+} (B_+ - B_-)}{\Gamma} dW_t.$$

On introduit le temps d'arrêt T_+^* :

$$T_+^* = \inf\{t \geq 0 : \zeta_t \geq l_2 \text{ ou } \zeta_t \leq -l_1\}.$$

On définit la probabilité $p_+(x) = P_{\zeta_0=x}(\{\zeta_{T_+^*} = -l_1\} | A_+)$. En utilisant la formule de Dynkin, on obtient l'équation différentielle ordinaire suivante:

$$\begin{cases} p_+'' + p_+' = 0 \\ p_+(-l_1) = 1, p_+(l_2) = 0. \end{cases}$$

La résolution de cette équation nous donne la probabilité de "refuser A_+ à tort", $p_+ = p_+(0)$ (cf. [Fleming-Rishel][chapV.th7.1]):

$$p_+ = \frac{1 - e^{-l_2}}{e^{l_1} - e^{-l_2}}.$$

Par un raisonnement identique sur A_- , on obtient l'expression de la probabilité d'erreur $p_- = p_-(0)$:

$$p_- = \frac{1 - e^{-l_1}}{e^{l_2} - e^{-l_1}}.$$

De même, on calcule les temps moyens pour prendre une décision en résolvant deux équations différentielles ordinaires (cf. [Fleming-Rishel][chapV. th7.1]). Plus précisément $E(T_+^*)$ est donné par la solution de l'équation

$$\begin{cases} \frac{\Lambda_+(B_+ - B_-)^2}{2\Gamma^2} g_+'' + \frac{\Lambda_+(B_+ - B_-)^2}{2\Gamma^2} g_+' + 1 = 0 \\ g_+(-l_1) = g_+(l_2) = 0 \end{cases}$$

au point 0.

On obtient $E(T_-^*)$ de façon similaire:

$$\begin{aligned} E(T_+^*) &= \frac{-2B_+(2 + \Delta t B_+)}{(B_+ - B_-)^2} [l_2 - p_+(l_1 + l_2)], \\ E(T_-^*) &= \frac{-2B_-(2 + \Delta t B_-)}{(B_+ - B_-)^2} [l_1 - p_-(l_1 + l_2)]. \end{aligned}$$

Remarque 5.3 Dans les applications on prend généralement $p_+ = p_- = \alpha$. Ainsi un simple calcul nous donne

$$l_1 = l_2 = \ln \frac{1 - \alpha}{\alpha}$$

et on note que

$$E(T_-^*) \approx \frac{B_-}{B_+} E(T_+^*).$$

Exemple 5.4 Pour $B_- = -1$, $B_+ = -0.25$ et $\alpha = 0.05$, on obtient $l_1 = l_2 = 2.94$. Les temps moyens sont alors $E(T_+^*) = 4.70$ et $E(T_-^*) = 18.79$.

Remarque 5.5 Pour comparer les temps moyens des deux procédures, on s'intéresse au signe de l'expression:

$$2\lambda^2 - (1 - 2\alpha) \ln \frac{1 - \alpha}{\alpha},$$

avec λ tel que $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} e^{-x^2/2} dx = \alpha$.

On vérifie numériquement que cette expression est positive, $\forall 0 < \alpha < 1/2$. On en déduit que pour une probabilité d'erreur donnée, les temps moyens de la première procédure sont plus grands que ceux de la deuxième. Cette différence est d'autant plus grande que α est petit.

En préparation de l'étude du cas $\varepsilon > 0$, nous réécrivons la statistique L_n sous une autre forme. En utilisant l'égalité

$$\tilde{y}_k (\tilde{y}_{k+1} - \tilde{y}_k) = -\frac{1}{2} ((\tilde{y}_{k+1} - \tilde{y}_k)^2 - \tilde{y}_{k+1}^2 + \tilde{y}_k^2)$$

on obtient

$$\begin{aligned} L_n &= \frac{B_+ - B_-}{2\Gamma^2} (\tilde{y}_{i_0+n+1}^2 - \tilde{y}_{i_0+1}^2) - \frac{B_+ - B_-}{2\Gamma^2} \sum_{k=i_0+1}^{i_0+n} (\tilde{y}_{k+1} - \tilde{y}_k)^2 \\ &\quad - \Delta t \frac{B_+^2 - B_-^2}{2\Gamma^2} \sum_{k=i_0+1}^{i_0+n} \tilde{y}_k^2. \end{aligned}$$

Lemme 5.6 Sur $A_+ \cup A_-$, on a l'estimation:

$$\sum_{k=i_0+1}^{i_0+n} (\tilde{y}_{k+1} - \tilde{y}_k)^2 - \Gamma^2 n \Delta t = \mathcal{O}(\sqrt{\Delta t}).$$

Preuve Démontrons, par exemple, ce résultat sur l'évènement A_+ .

Sur A_+ , on a :

$$\tilde{y}_{k+1} = \tilde{y}_k + \Delta t H_+ B_+ x_k + \sqrt{\Delta t} \sigma_+ H_+ u_k,$$

d'où l'expression:

$$\begin{aligned} \sum_{k=i_0+1}^{i_0+n} (\tilde{y}_{k+1} - \tilde{y}_k)^2 &= \Delta t^2 H_+^2 B_+^2 \sum_{k=i_0+1}^{i_0+n} x_k^2 + \Delta t \Gamma^2 \sum_{k=i_0+1}^{i_0+n} u_k^2 \\ &\quad + 2\Delta t^{3/2} H_+^2 \sigma_+ B_+ \sum_{k=i_0+1}^{i_0+n} x_k u_k. \end{aligned}$$

On obtient l'estimation:

$$\mathbf{E} \left[\sum_{k=i_0+1}^{i_0+n} (\tilde{y}_{k+1} - \tilde{y}_k)^2; A_+ \right] \leq \Delta t^2 H_+^2 B_+^2 \sum_{k=i_0+1}^{i_0+n} \mathbf{E}(x_k^2) + \Delta t n \Gamma^2.$$

L'égalité

$$\mathbf{E} \left[\left(\sum_{k=i_0+1}^{i_0+n} u_k^2 \right) \left(\sum_{k=i_0+1}^{i_0+n} x_k u_k \right) \right] = 0,$$

entraîne

$$\begin{aligned}
& \mathbf{E} \left[\left(\sum_{k=i_0+1}^{i_0+n} (\tilde{y}_{k+1} - \tilde{y}_k)^2 - \Gamma^2 n \Delta t \right)^2 ; A_+ \right] \\
& \leq \mathbf{E} \left[\left(\Delta t^2 H_+^2 B_+^2 \sum_{k=i_0+1}^{i_0+n} x_k^2 + 2 \Delta t^{3/2} H_+^2 \sigma_+ B_+ \sum_{k=i_0+1}^{i_0+n} x_k u_k \right)^2 \right] \\
& \quad + \Delta t^2 \Gamma^4 \mathbf{E} \left[\left(\sum_{k=i_0+1}^{i_0+n} u_k^2 \right)^2 \right] + 2 \Delta t^3 \Gamma^2 H_+^2 B_+^2 \mathbf{E} \left[\left(\sum_{k=i_0+1}^{i_0+n} u_k^2 \right) \left(\sum_{k=i_0+1}^{i_0+n} x_k^2 \right) \right] \\
& \quad - \Delta t^2 n^2 \Gamma^4 - 2 \Delta t^3 n \Gamma^2 H_+^2 B_+^2 \sum_{k=i_0+1}^{i_0+n} \mathbf{E}(x_k^2). \tag{24}
\end{aligned}$$

On remarque que

$$\mathbf{E} \left[\left(\sum_{k=i_0+1}^{i_0+n} u_k^2 \right)^2 \right] = \sum_{k=i_0+1}^{i_0+n} \mathbf{E}(u_k^4) + n(n-1) = n(n+2),$$

puisque $\{u_k\}$ est un bruit blanc standard. Ainsi on a:

$$\Delta t^2 \Gamma^4 \mathbf{E} \left[\left(\sum_{k=i_0+1}^{i_0+n} u_k^2 \right)^2 \right] - \Delta t^2 n^2 \Gamma^4 = 2 \Delta t^2 n \Gamma^4$$

On utilise l'inégalité $(a+b)^2 \leq 2a^2 + 2b^2$ pour traiter le premier terme de (24) et l'inégalité de Cauchy - Schwarz pour le troisième. Vu que les moments d'ordre 2 et 4 de x_k sont finis sur l'intervalle $[a, b]$ et que $n \leq (b-a)/\Delta t$, on obtient la majoration suivante:

$$\mathbf{E} \left[\left(\sum_{k=i_0+1}^{i_0+n} (\tilde{y}_{k+1} - \tilde{y}_k)^2 - \Gamma^2 n \Delta t \right)^2 \right] \leq c_{a,b} \Delta t,$$

où $c_{a,b}$ est une constante positive dépendante de l'intervalle $[a, b]$ fixé.

Sur l'événement A_- , la démonstration est similaire. ■

5.1.2 Le cas $\varepsilon > 0$

On définit la statistique L_n^ε comme il suit:

$$L_n^\varepsilon = \frac{B_+ - B_-}{2 \Gamma^2} (y_{i_0+n+1}^2 - y_{i_0+1}^2) - \frac{1}{2} n \varepsilon (B_+ - B_-)$$

$$-\varepsilon \frac{B_+^2 - B_-^2}{2\Gamma^2} \sum_{k=i_0+1}^{i_0+n} y_k^2. \quad (25)$$

On considère $\Delta t = \varepsilon$ et on s'intéresse à l'écart $L_n^\varepsilon - L_n$.

Théorème 5.7 *Sur les événements \mathcal{C} et $\hat{\mathcal{C}}$ définis dans la section 3, on a l'estimation*

$$L_n^\varepsilon - L_n = \mathcal{O}(\sqrt{\varepsilon}).$$

Preuve

On démontre d'abord que $L_n^\varepsilon - L_n$ est d'ordre $\mathcal{O}(\sqrt{\varepsilon})$ sur $A_+ \cup A_-$.

$$\begin{aligned} L_n^\varepsilon - L_n &= \frac{B_+ - B_-}{2\Gamma^2} \left((y_{i_0+n}^2 - \tilde{y}_{i_0+n}^2) - (y_{i_0+1}^2 - \tilde{y}_{i_0+1}^2) \right) \\ &\quad + \frac{B_+ - B_-}{2\Gamma^2} \left[\sum_{k=i_0+1}^{i_0+n} (\tilde{y}_{k+1} - \tilde{y}_k)^2 - \Gamma^2 n \varepsilon \right] \\ &\quad - \varepsilon \frac{B_+^2 - B_-^2}{2\Gamma^2} \sum_{k=i_0+1}^{i_0+n} (y_k^2 - \tilde{y}_k^2). \end{aligned}$$

L'égalité

$$y_k^2 - \tilde{y}_k^2 = \sqrt{\varepsilon} v_k (2h(x_k) + \sqrt{\varepsilon} v_k)$$

implique

$$\begin{aligned} L_n^\varepsilon - L_n &= \sqrt{\varepsilon} \frac{B_+ - B_-}{\Gamma^2} (v_{i_0+n+1} h(x_{i_0+n+1}) - v_{i_0+1} h(x_{i_0+1})) \\ &\quad + \frac{B_+ - B_-}{2\Gamma^2} \left[\sum_{k=i_0+1}^{i_0+n} (h(x_{k+1}) - h(x_k))^2 - \Gamma^2 n \varepsilon \right] \\ &\quad + \varepsilon \frac{B_+ - B_-}{2\Gamma^2} (v_{i_0+n+1}^2 + v_{i_0+1}^2) - \varepsilon^{3/2} \frac{B_+^2 - B_-^2}{2\Gamma^2} \sum_{k=i_0+1}^{i_0+n} v_k h(x_k) \\ &\quad + \varepsilon^2 \frac{B_+^2 - B_-^2}{2\Gamma^2} \sum_{k=i_0+1}^{i_0+n} v_k^2. \end{aligned}$$

Sachant que $\{v_k\}$ est un bruit blanc gaussien standard et que tous les moments de x_k sont finis pour $k \leq m$, on déduit à l'aide du lemme 5.6 que $L_n^\varepsilon - L_n$ est d'ordre $\mathcal{O}(\sqrt{\varepsilon})$ sur $A_+ \cup A_-$.

Pour raisonner sur \mathcal{C} , on remarque tout simplement qu'on a la majoration

$$\begin{aligned} \mathbf{E}[(L_n^\varepsilon - L_n)^2 ; \mathcal{C}] &\leq \mathbf{E}[(L_n^\varepsilon - L_n)^2 ; A_+ \cup A_-] \\ &\quad + \mathbf{E}[(L_n^\varepsilon - L_n)^2 ; \mathcal{C} \cap (A_+ \cup A_-)^c]. \end{aligned}$$

De l'inégalité de Cauchy-Schwarz, on obtient

$$\mathbf{E}[(L_n^\varepsilon - L_n)^2 ; \mathcal{C}] \leq \mathbf{E}[(L_n^\varepsilon - L_n)^4]^{1/2} P[\mathcal{C} \cap (A_+ \cup A_-)^c]^{1/2} + C\varepsilon.$$

Du fait que $\mathbf{E}[(L_n^\varepsilon - L_n)^4] \leq C$ et de la proposition 3.1, on déduit le résultat sur l'évènement \mathcal{C} .

De même, du théorème 3.6, on déduit le résultat sur l'évènement $\hat{\mathcal{C}}$. ■

Etant donné que les bornes déterminées dans le paragraphe 5.1.1 sont d'ordre $\mathcal{O}(1)$, il est donc légitime, d'après le résultat précédent, de les utiliser pour tester L_n^ε .

D'après la remarque 5.5, nous avons choisi d'implémenter le test selon sa deuxième version.

5.2 Test du rapport de vraisemblance sur les sorties des filtres de Kalman - Bucy

On montre que la statistique \hat{L}_n , utilisée dans le paragraphe 4.2 sous l'hypothèse (HD1), permet encore de décider entre les alternatives " A_+ " et " A_- " sous l'hypothèse (HD2).

Soit $i_0 < i_1 < i_0 + m$. On a

$$\hat{L}_n = \frac{1}{\varepsilon} \sum_{k=i_1}^{i_1+n} Z_k y_{k+1} - \frac{1}{2\varepsilon} \sum_{k=i_1}^{i_1+n} \left((H_+(1 + \varepsilon B_+) \hat{x}_k^+)^2 - (H_-(1 + \varepsilon B_-) \hat{x}_k^-)^2 \right),$$

où $Z_k = H_+(1 + \varepsilon B_+) \hat{x}_k^+ - H_-(1 + \varepsilon B_-) \hat{x}_k^-$.

On rappelle la définition du terme de précision du test:

$$\hat{R}_n = \frac{1}{\varepsilon} \sum_{k=i_1}^{i_1+n} Z_k^2.$$

Sur A_+ , à partir de l'expression récursive de Z_k (équation (16)), on obtient:

$$\begin{aligned} \mathbf{E}[Z_{k+1}^2 ; A_+] &\leq (1 + \varepsilon B_-)^2 (1 - H_- K_-)^2 \mathbf{E}[Z_k^2 ; A_+] \\ &\quad + \varepsilon^2 (1 + \varepsilon B_+)^2 (B_+ - B_-)^2 H_+^2 \mathbf{E}[\hat{x}^{+2} ; A_+] \\ &\quad + \varepsilon [H_+ K_+ (1 + \varepsilon B_+) - H_- K_- (1 + \varepsilon B_-)]^2 \vartheta^+ \\ &\quad + 2\varepsilon (1 + \varepsilon B_+) (1 + \varepsilon B_-) |B_+ - B_-| (1 - H_- K_-) H_+ \mathbf{E}[Z_k \hat{x}_k^+ ; A_+]. \end{aligned}$$

On utilise l'inégalité

$$\begin{aligned} \mathbb{E}[Z_k \hat{x}_k^+; A_+] &\leq \frac{H_- K_-}{2\varepsilon(1+\varepsilon B_-)(1+\varepsilon B_+)|B_+ - B_-|H_+} \mathbb{E}[Z_k^2; A_+] \\ &\quad + \frac{\varepsilon(1+\varepsilon B_-)(1+\varepsilon B_+)|B_+ - B_-|H_+}{2H_- K_-} \mathbb{E}[\hat{x}^{+2}; A_+]. \end{aligned}$$

Sous (HD2), la différence $H_+ K_+ - H_- K_-$ est d'ordre $\mathcal{O}(\varepsilon)$.

On utilise le fait que

$$0 < (1 + \varepsilon B_+)^2 (1 - H_- K_-)^2 + H_- K_- (1 - H_- K_-) < 1 - H_- K_- + C\varepsilon < 1.$$

Sous l'hypothèse (H4), on obtient, de la même façon que dans le paragraphe 4.2,

$$\mathbb{E}[Z^2; A_+] \leq \Lambda_+ \varepsilon^2 + C\varepsilon^3,$$

où

$$\hat{\Lambda}_+ = \frac{(B_+ - B_-)^2}{|B_+|(2 + \varepsilon B_+)} \vartheta_+.$$

Sur A_- , on fait un raisonnement similaire.

Remarque 5.8 Le processus $\{Z_k\}$ est d'ordre $\mathcal{O}(\sqrt{\varepsilon})$ sous l'hypothèse (HD1) et d'ordre $\mathcal{O}(\varepsilon)$ sous l'hypothèse (HD2).

De même que pour le test précédent, on considère deux types de procédures.

Première procédure:

Pour r fixé, on introduit le temps d'arrêt

$$N^* = \inf\{n \geq 1 : R_n \geq r\} \wedge m.$$

On calcule \hat{L}_n jusqu'à ce qu'on ait $R_n \geq r$. Par un raisonnement similaire à celui du paragraphe 5.1, on choisit $r = \max(r_-, r_+)$ avec

$$r_- = 4\lambda^2 \vartheta^- \text{ et } r_+ = 4\lambda^2 \vartheta^+.$$

La constante λ est obtenue par inversion de la fonction de répartition de la loi $\mathcal{N}(0, 1)$ calculée pour la probabilité α . Les estimations des temps moyens pour prendre une décision sont données par

$$\begin{aligned} \varepsilon N_-^* &= 4\lambda^2 \frac{-B_-(2 + \varepsilon B_-)}{(B_+ - B_-)^2}, \\ \varepsilon N_+^* &= 4\lambda^2 \frac{-B_+(2 + \varepsilon B_+)}{(B_+ - B_-)^2}. \end{aligned}$$

Remarque 5.9 Les temps moyens pour prendre une décision associés à cette procédure sont identiques à ceux du test précédent quand on considère le même type d'approche (cf. paragraphe 5.1).

Deuxième procédure:

Soient $l_1 > 0$ et $l_2 > 0$ des bornes fixées.

On introduit le temps d'arrêt N^* :

$$N^* = \inf\{n > i_1 : \hat{L}_n \geq l_2 \text{ ou } \hat{L}_n \leq -l_1\} \wedge m.$$

Raisonnons sur l'événement A_+ , par exemple.

Nous savons, d'après ce qui précède, que sous l'hypothèse (H_4) le processus $\xi_k = Z_k \nu_{k+1}$ est approximativement un processus mélangeant. Par un raisonnement similaire à celui du paragraphe 4.2, pour $t = \varepsilon n$, on approche \hat{L}_n par le processus de diffusion $\{\zeta_t\}$ défini par:

$$d\zeta_t = \hat{\Lambda}_+ / 2 dt + \sqrt{\vartheta^+ \hat{\Lambda}_+} dW_t^+.$$

On obtient la probabilité d'erreur p_+ , "probabilité de refuser A_+ à tort":

$$p_+ = \frac{1 - e^{-\eta_+ l_2}}{e^{\eta_+ l_1} - e^{-\eta_+ l_2}},$$

où $\eta_+ = 1/\vartheta^+$ ainsi que le temps moyen pour prendre une décision sur A_+ :

$$E(T_+^*) = 2 \left(\frac{l_2}{\hat{\Lambda}_+} - p_+ \frac{(l_1 + l_2)}{\hat{\Lambda}_+} \right).$$

De même, on obtient la probabilité d'erreur p_- , "probabilité de refuser A_- à tort":

$$p_- = \frac{1 - e^{-\eta_- l_1}}{e^{\eta_- l_2} - e^{-\eta_- l_1}},$$

où $\eta_- = 1/\vartheta^-$ et le temps moyen pour prendre une décision sur A_- :

$$E(T_-^*) = 2 \left(\frac{l_1}{\hat{\Lambda}_-} - p_- \frac{(l_1 + l_2)}{\hat{\Lambda}_-} \right).$$

Remarque 5.10 Sous (HD_1) les temps d'attente moyens sont d'ordre $\mathcal{O}(\varepsilon)$ tandis que sous (HD_2) ils sont d'ordre $\mathcal{O}(1)$. On peut s'attendre donc à ce que, dans ce deuxième cas, il y ait des intervalles où aucune décision ne sera prise.

Remarque 5.11 Sous l'hypothèse (HD2), l'écart $\vartheta^+ - \vartheta^-$ est d'ordre $\mathcal{O}(\varepsilon)$, donc les bornes l_1 et l_2 sont "très proches". On a une situation similaire à celle décrite dans la remarque 5.5, i.e. les temps moyens d'attente sont plus longs dans la première procédure que dans la deuxième.

Le test du rapport de vraisemblance sur les sorties des filtres de Kalman sera mis en œuvre selon la deuxième procédure.

6 Application

Nous rappelons que notre objectif est d'étudier le comportement des filtres proposés par rapport à la solution optimale du problème de filtrage et de comparer les différentes procédures permettant de les obtenir. Nous devons donc, en premier lieu, définir des critères de comparaison.

6.1 Les critères de comparaison

Une première idée est de considérer, pour une probabilité d'erreur fixée,

- l'amplitude moyenne des intervalles détectés, pour comparer les tests de détection
- le temps moyen nécessaire pour prendre une décision, pour comparer les tests de détermination du signe.

Pour comparer les tests en termes de probabilité d'erreur, il semble naturel de s'intéresser aux pourcentages suivants:

- pourcentage d'intervalles détectés corrects.
- pourcentage de décisions correctes (i.e. nombre de décisions correctes par rapport au nombre de décisions prises).

Nous qualifions une décision de correcte lorsque le signe choisi correspond à celui de la trajectoire. ceci sur un intervalle de monotonie correct.

Les temps moyens pour atteindre une décision sont calculés d'abord théoriquement (voir les paragraphes 4 et 5) et sont ensuite confrontés aux résultats obtenus par simulation.

6.2 Exemples

L'étude du problème sous l'hypothèse ($HD2$) étant notre principal centre d'intérêt, nous considérons des exemples vérifiant cette hypothèse et nous appliquons les différents tests décrits dans le paragraphe 5. Nous étudions l'influence de la fonction d'observation sur l'efficacité des tests, dans l'exemple 6.2 et celle de la dérive dans l'exemple 6.3. Nous nous intéressons ensuite à un exemple vérifiant l'hypothèse ($HD1$), présenté dans [FJSZ]. Ceci nous permettra d'illustrer que dans ce cas, les décisions sont prises beaucoup plus rapidement que sous ($HD2$). Nous justifierons ce fait par l'étude du comportement de la solution de l'équation de Zakai.

Dans le paragraphe 5, sous (HD2), nous avons proposé une méthode de détermination des bornes pour les tests du signe et d'estimation des temps d'attente sous l'hypothèse (H4): $B_- < 0$ et $B_+ < 0$. Il est intéressant d'étudier les résultats obtenus dans le cas où l'hypothèse (H4) n'est pas vérifiée. C'est ce que nous nous proposons de faire dans l'exemple 6.6.

Pour chaque exemple, nous commentons les résultats statistiques sur le comportement des tests. On peut consulter ces résultats dans l'annexe A.

Remarque 6.1 On rappelle que la détermination des bornes utilisées dans les tests de décision sur le signe repose sur la possibilité d'atteindre une situation stationnaire. Une fois les passages à zéro détectés, nous devons attendre un "certain" temps, n_0 , avant de cumuler les statistiques de décision des tests. Ce temps est cependant assez court puisque, comme nous l'avons déjà remarqué, à l'intérieur d'un intervalle de monotonie, le système a tendance à accompagner rapidement l'évolution du système linéaire associé. Nous avons choisi pour ces exemples $n_0 = 6$.

6.2.1 Exemples sous l'hypothèse (HD2)

Exemple 6.1 On considère le système

$$\begin{cases} x_{k+1} = x_k + 0.01 b(x_k) + \sqrt{0.01} u_k, & x_0 \sim \mathcal{N}(-5, 0.1) \\ y_k = |x_k| + \sqrt{0.01} v_k, \end{cases}$$

avec

$$b(x) = \begin{cases} -x, & x \leq 0 \\ -0.25x, & x \geq 0. \end{cases}$$

On s'intéresse tout d'abord à une trajectoire simulée sur l'intervalle de temps $[0, 10]$.

Dans ce cas, les résultats des différentes procédures étant similaires, nous illustrons par la figure 1, uniquement ceux des tests de détection et du rapport de vraisemblance sur les observations.

Dans cette figure, la trajectoire du signal à estimer est tracée en trait plein et le filtre proposé en pointillé. Les intervalles de monotonie détectés sont signalés en noir dans le cas où une décision est prise sur le signe et en gris s'il n'y a pas de décision. Nous rappelons que sur un intervalle de monotonie, si le test appliqué décide "signe positif", (respectivement "signe négatif") on trace la sortie du filtre de Kalman-Bucy (7) (respectivement la sortie du filtre (8)). Si le test ne prend pas de décision, les sorties des deux filtres sont représentées.

Il est intéressant de regarder le comportement de la solution optimale pour ce problème. Dans la figure 2 on représente, en plus de la trajectoire (en trait plein), l'estimation optimale obtenue par résolution numérique de l'équation de Zakai (en trait pointillé), entourée par sa région de confiance correspondante à une probabilité de 95%. Un générateur automatique de programmes FORTRAN a été utilisé pour obtenir cette solution.

Dans les figures 3 à 6, on trace la densité conditionnelle à divers instants de l'intervalle $[0, 10]$. Le trait vertical correspond à la valeur du signal simulé à l'instant t . On note qu'au moment où la décision sur le signe est prise, la loi conditionnelle approche une loi gaussienne et que près d'un passage à zéro, la loi conditionnelle a deux "bosses".

On peut constater, par exemple, qu'à l'instant $t = 6.5$ les tests, ayant déjà détecté un intervalle de monotonie, ne sont pas encore capables de décider sur le signe. Cette décision sera prise aux environs de $t = 8.48$. En regardant le comportement de la solution de l'équation de Zakai, on remarque l'existence de deux "bosses" à partir de l'instant $t = 3$, approximativement. A l'instant $t = 8.48$, instant où les tests de décision sur le signe sont capables de prendre une décision, une des "bosses" de la densité conditionnelle est déjà négligeable (voir la figure 5), ce qui explique le fait qu'une décision puisse être prise. A cet instant, le filtre proposé et l'estimation optimale ont un comportement similaire.

On s'intéresse maintenant aux résultats numériques obtenus pour une trajectoire simulée sur l'intervalle de temps $[0, 100]$ (voir les tableaux 1 et 2).

On constate que

- Le test de détection sur les sorties du filtre de Kalman-Bucy permet de détecter des intervalles de monotonie plus longs que le test sur les observations; cependant l'erreur commise est aussi plus grande.
- Le fait d'augmenter la borne de détection (par exemple en diminuant la probabilité d'erreur α) entraîne évidemment la détection d'intervalles de monotonie plus petits. Il semble que nous contrôlons moins bien la probabilité d'erreur du test sur la sortie des filtres de Kalman-Bucy que celle du test sur les observations. Voyons par exemple ce qui se passe quand on prend $\alpha_d = 5\%$ (voir le tableau 2). Le pourcentage d'erreur de détection p_2 est supérieur à celui qui avait été imposé.
- Les résultats des tests de décision sur le signe basés sur le rapport de vraisemblance sur les observations ou sur les sorties du filtre de Kalman sont comparables. Toutefois le temps moyen pour prendre une décision semble être légèrement plus petit dans le premier cas que dans le deuxième.

- Le fait de diminuer la probabilité d'erreur pour la décision sur le signe entraîne naturellement une réduction du pourcentage d'intervalles avec décision et des temps d'attente plus grands pour prendre ces décisions.

Remarque 6.1 Les temps moyens pour prendre une décision calculés théoriquement sont beaucoup plus élevés que ceux vérifiés en pratique.

Exemple 6.2 On modifie la fonction d'observation dans l'exemple 6.1. On considère :

$$h(x) = 0.4 |x|.$$

On obtient les résultats présentés dans le tableau 3.

Le fait de diminuer le coefficient $|H_+| = |H_-|$ semble avoir pour effet l'augmentation des pourcentages de détection d'intervalles corrects. Autrement dit, lorsqu'on diminue le coefficient $|H_+| = |H_-|$, les constantes des tests de détection augmentent. On détecte plus de passages à zéro et les longueurs des intervalles de monotonie diminuent. Ce comportement des tests est satisfaisant vu qu'une diminution du coefficient $|H_+| = |H_-|$ est équivalente à une augmentation du bruit d'observation.

On vérifie également une "legère" diminution des temps d'attente pour prendre une décision. Ceci s'explique du fait que les statistiques des tests du signe sont calculés loin des passages à zéro.

Dans ce cas, le test de détection des passages à zéro basé sur les sorties d'un des filtres de Kalman-Bucy est meilleur que celui basé sur les observations.

Exemple 6.3 On considère les mêmes paramètres que dans l'exemple précédent sauf pour la dérivée $b(x)$:

$$b(x) = \begin{cases} -5x, & x \leq 0 \\ -0.25x, & x \geq 0. \end{cases}$$

Les résultats sont à consulter dans le tableau 4.

On remarque que le pourcentage de décisions est maintenant plus élevé et que le temps moyen pour prendre une décision diminue considérablement. Ceci illustre parfaitement les calculs théoriques présentés dans le paragraphe 5. Le fait de diminuer le rapport $|B_+|/(B_+ - B_-)^2$ (respectivement $|B_-|/(B_+ - B_-)^2$) a pour effet d'augmenter le terme de précision dans les statistiques L et \hat{L} . On obtient donc de meilleurs résultats, notamment pour le temps d'attente "côté positif" (respectivement "côté négatif").

6.2.2 Exemple sous l'hypothèse (HD1)

Exemple 6.4 Nous changeons la fonction d'observation de l'exemple 6.1:

$$h(x) = \begin{cases} -x, & x \leq 0, \\ 2x, & x \geq 0. \end{cases}$$

Le système considéré vérifie l'hypothèse (HD1).

Par les figures 7 à 10, nous illustrons les résultats des différentes procédures appliquées à une trajectoire simulée sur l'intervalle de temps $[0, 10]$. Nous remarquons que le test de la variation quadratique décide plus souvent que le test sur les sorties des filtres de Kalman-Bucy.

Le comportement de l'estimation optimale et sa région de confiance à 95% représentés dans la figure 11, montrent clairement que dans ce cas, la loi conditionnelle approche beaucoup plus rapidement une loi gaussienne après un passage à zéro que dans l'exemple 6.1 (voir la figure 2). Ceci justifie le fait que nous décidons du signe de x_k plus souvent sous l'hypothèse (HD1) que sous (HD2).

La comparaison entre les figures 12 à 14 et les figures 3 à 6 (voir l'exemple 6.1) illustre à nouveau la remarque précédente. Dans ce cas, une des deux "bosses" de la loi conditionnelle s'estompe rapidement aux environs des instants de décision ($t = 6.86$ et $t = 6.88$), tandis que dans l'exemple 6.1 on ne retrouve un tel comportement qu'aux alentours de l'instant $t = 10$.

Les résultats numériques des tests appliqués à une trajectoire simulée sur l'intervalle de temps $[0, 100]$ (voir le tableau 5) mettent également en évidence que les temps d'attente pour prendre une décision sous l'hypothèse (HD1) sont nettement plus courts que sous l'hypothèse (HD2).

Remarque 6.5 Quelques remarques sur les résultats des procédures appliquées à plusieurs trajectoires simulées sur un intervalle de temps $[0, 100]$.

- Les tests détectent moins de passages à zéro lorsque l'écart $|H_+ - H_-|$ augmente et le taux d'erreur est supérieur au taux souhaité. De façon plus nette que sous (HD2), nous contrôlons "moins bien" la probabilité d'erreur du test de détection sur la sortie d'un des filtres de Kalman-Bucy que celle du test sur les observations.
- Pour une différence $|\sigma_+ H_+ - \sigma_- H_-| = 0.5$, les tests de détermination du signe ont des taux d'erreur supérieurs à ceux demandés. Le test de la variation quadratique a tendance à décider plus souvent que le test du rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy, mais semble généralement moins fiable. (cf. [FJSZ]).

- En augmentant l'écart $|\sigma_+ H_+ - \sigma_- H_-|$, les tests de détermination du signe décident plus souvent et prennent moins de fausses décisions.

6.2.3 Un exemple sous l'hypothèse (HD2) ne vérifiant pas (H4)

Exemple 6.6 On considère le système suivant:

$$\begin{cases} x_{k+1} = x_k + 0.01 b(x_k) + \sqrt{0.01} u_k, & x_0 \sim \mathcal{N}(-5, 0.1) \\ y_k = |x_k| + \sqrt{0.01} v_k, \end{cases}$$

avec

$$b(x) = \begin{cases} -x & x \leq 0 \\ 0.2x & x \geq 0 \end{cases}$$

Remarque 6.7 Le coefficient de dérive B_+ étant positif et B_- étant négatif, il y a "peu" de passages à zéro et le signal a tendance à "fuir" vers l'infini dès qu'il est du côté positif.

Nous étudions un cas où l'hypothèse (HD2) est vérifiée mais pas (H4). Cependant on applique les tests présentés dans la section 5. On rappelle que les bornes utilisées dans ces tests ont été déterminées sous l'hypothèse (H4) et ne sont plus justifiées dans cet exemple. Il est donc intéressant de regarder les résultats qu'elles engendrent dans une telle situation.

Dans la figure 15, nous traçons le filtre approché construit à partir du test de détection et du test de rapport de vraisemblance basés sur les observations. Les résultats obtenus par les autres procédures sont similaires.

En comparant le filtre proposé et l'estimation optimale (voir la figure 16), on remarque que leurs comportements sont semblables sur l'intervalle de monotonie détecté $[0, 3]$ et aussi à partir de l'instant $t = 7$, environ. Ceci est satisfaisant vu que sur l'intervalle de monotonie $[5.80, 10]$, les instants de décision sur le signe sont $t = 7.15$, pour le test sur les observations, et $t = 7.20$ pour le test sur les sorties des filtres de Kalman-Bucy.

On se permet de conclure qu'au moins pour certains cas ne vérifiant pas l'hypothèse (H4), on peut encore utiliser les bornes calculées sous cette hypothèse et obtenir des résultats "corrects".

Remarque 6.8 Dans cette situation où l'hypothèse (H4) n'est pas vérifiée, nous ne savons pas estimer théoriquement le temps d'attente d'une décision.

7 Conclusion

On a décrit des tests permettant de déterminer les intervalles de linéarité de la fonction d'observation, pour un système linéaire par morceaux en temps discret avec petit bruit d'observation. Deux types de tests sont nécessaires: un premier test permet de détecter les intervalles où il n'y a pas de passages à zéro; un deuxième test permet de décider du signe pris par la trajectoire sur cet intervalle. Sous l'hypothèse (HD2), différents tests ont été proposés. Pour la détection de l'intervalle de monotonie, on a le choix entre un test basé sur les observations et un test basé sur les sorties d'un des filtres de Kalman-Bucy. Pour l'étape de décision du signe, on décrit un test de rapport de vraisemblance sur les observations et un autre test de rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy.

L'efficacité des tests de détection dépend de la détermination de la valeur d'une borne, le "cutoff". Une formule permettant d'obtenir cette valeur a été justifiée théoriquement pour le test basé sur les observations. Ceci n'a pas été de même pour le test basé sur les sorties du filtre de Kalman-Bucy. Il nous semble donc naturel que ce deuxième test se comporte moins bien que le premier. Quant aux tests de détermination du signe, on a vérifié que leur comportement est semblable. Des formules permettant de déterminer les bornes et les temps moyens d'attente pour prendre une décision ont été obtenues théoriquement sous l'hypothèse (H4). Les probabilités d'erreur fixées sont à comparer avec celles obtenues sur le problème simulé ainsi que les temps moyens d'attente. On remarque de manière générale, que les temps moyens d'attente calculés théoriquement sont plus longs que ceux obtenus en pratique. On note également que l'ordre de grandeur des temps moyens d'attente théoriques et empiriques est plus grand si on se place sous l'hypothèse (HD2) que sous l'hypothèse (HD1).

On peut envisager d'étendre l'étude ici présentée à des systèmes où les fonctions b , σ et h possèdent l points critiques, avec $l \geq 1$. On peut aussi considérer des systèmes de dimension $n \geq 1$, avec des bruits corrélés.

Une extension au cas de fonctions monotones par morceaux existe dans [Fleming-Pardoux], sous l'hypothèse (HD1). Des essais ont été faits par [FJSZ] dans le cas d'une dynamique non linéaire. Il reste à justifier son analogue sous l'hypothèse (HD2).

Références

- [Bensoussan] A. Bensoussan : *On some approximation techniques in non-linear filtering*, Stoch. Diff. Syst., Stoch. Control and Appl., W.H. Fleming & P.L. Lions eds., IMA Vol 10, Springer, 1987.
- [Billingsley] Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, New York, 1968.
- [DiMasi-Runggaldier] G. DiMasi and W. Runggaldier, *On measure transformation for combined filtering and parameter estimation in discrete time*, Systems & Control Letters 2, pp 57-62, 1982.
- [Either-Kurtz] S.N. Either and T.G. Kurtz : *Markov Processes: Characterization and Convergence*, John Wiley & Sons, 1986.
- [FJP] W.H. Fleming, D. Ji and E. Pardoux : *Piecewise linear filtering with small observation noise*, Proc. 8th Int. Conference on Analysis and Optimization of Systems, Antibes 1988, Lect. Notes in Control and Info. Sci., n° 111, (pp 725-739), Springer, 1988.
- [FJSZ] W.H. Fleming, D. Ji, P. Salame and Q. Zhang : *Discrete Time Piecewise Linear Filtering with Small Observation Noise*, Brown University Providence RI 02912, Division of Applied Mathematics, September 1988 LCDS/CCS 88-27.
- [Fleming-Pardoux] W.H. Fleming, E. Pardoux : *Piecewise monotone filtering with small observation noise*, Siam J. Control, Vol. 27, No. 5, pp 1156-1181, Sep. 1989.
- [Fleming-Rishel] W. H. Fleming and R.W. Rishel : *Deterministic and Stochastic Optimal Control*, Tome 1, Applications of Mathematics, Springer - Verlag, 1975.
- [Ji] D. Ji. *Nonlinear filtering with small observation noise*, Ph. D. Thesis, Brown University, 1987.
- [KBS] R. Katzur, B.Z. Bobrovsky and Z. Schuss : *Asymptotic analysis of the optimal filtering problem for one-dimensional diffusions measured in a low noise channel*, SIAM J. Appl. Math. 44 (1984), Part I: 591-604, Part II: 1176-1191.

- [Le Gland] F. Le Gland : *Estimation de paramètres dans les processus stochastiques, en observation incomplète – Application à un problème de radio-astronomie*, Thèse présentée à l'Université Paris IX-Dauphine le 26 Novembre 1981 pour obtenir le titre de Docteur-Ingenieur.
- [Milheiro] P. Milheiro de Oliveira : *Filtres approchés pour un problème de filtrage non linéaire discret avec petit bruit d'observation*, RR 1142, INRIA, Dec. 1989.
- [Neveu] J. Neveu : *Martingales à temps discret*, Masson & Cie Edit., 1972.
- [Picard] J. Picard : *Nonlinear filtering of one-dimensional diffusions in the case of a high signal-to-noise ratio*. 1986 - SIAM J. Appl. Math. (46),6,1098-1125.
- [Picard] J. Picard : *Nonlinear filtering and smoothing with high signal-to-noise ratio*, Stochastic Processes in Physics and Engineering (Bielefeld 1986), Reidel 1988.
- [Picard] J. Picard : *Efficiency of the extended Kalman filter for nonlinear systems with small observation noise*, Rapport de Recherche INRIA n° 1068, 1989.
- [Roubaud] M.C. Roubaud : *Thèse à paraître*.

A Résultats numériques et représentations graphiques des exemples présentés dans la section 6

A.1 Résultats numériques

Nous introduisons les notations suivantes:

TDO: test de détection sur les observations.

TDK: test de détection sur la sortie d'un des filtres de Kalman-Bucy.

TLO: test du rapport de vraisemblance sur les observations.

TLK: test du rapport de vraisemblance sur les sorties des filtres de Kalman-Bucy.

α_d : niveau d'erreur pour le test de détection.

α_s : niveau d'erreur pour le test de détermination du signe.

C : "cutoff" du test de détection.

$-l_1, l_2$: intervalle de confiance du test de détermination du signe.

ET_-^*, ET_+^* : temps moyens théoriques d'attente d'une décision.

p_1 : pourcentage de pas de temps détectés.

p_2 : pourcentage d'intervalles de monotonie corrects.

p_3 : pourcentage de pas de temps avec décision.

p_4 : pourcentage de décisions correctes.

\bar{T}_-, \bar{T}_+ : moyennes empiriques des temps d'attente d'une décision.

Test de détection	TDO	TDO	TDK	TDK
Test du signe	TLO	TLK	TLO	TLK
C	0.143	0.143	0.143	0.143
$-l_1, l_2$	-2.94 2.94	-7.70 7.67	-2.94 2.94	-7.70 7.67
ET_-^*, ET_+^*	18.8 4.71	18.9 4.69	18.8 4.71	18.9 4.69
p_1	84.5%	84.5%	86.3%	86.3%
p_2	96.3%	96.3%	93.2%	93.2%
p_3	40.6%	40.6%	41.7%	41.7%
p_4	100%	100%	100%	100%
\bar{T}_-, \bar{T}_+	1.3 1.5	1.3 1.5	2.1 1.6	1.7 1.6

Tableau 1: résultats numériques pour l'exemple 6.1 ($\alpha_d = 5\%$, $\alpha_s = 5\%$)

Test de détection	TDO	TDO	TDK	TDK
Test du signe	TLO	TLK	TLO	TLK
C	0.17	0.17	0.17	0.17
$-l_1 \ l_2$	-2.94 2.94	-7.70 7.67	-2.94 2.94	-7.70 7.67
$ET_-^* \ ET_+^*$	18.8 4.71	18.9 4.69	18.8 4.71	18.9 4.69
p_1	82.4%	82.4%	84.2%	84.2%
p_2	97.9%	97.9%	97.5%	97.5%
p_3	39.0%	39.0%	41.4%	41.4%
p_4	100%	100%	100%	100%
$\bar{T}_- \ \bar{T}_+$	1.3 1.5	1.3 1.5	2.1 1.5	1.7 1.5

Tableau 2: résultats numériques pour l'exemple 6.1 ($\alpha_d = 2.5\%$, $\alpha_s = 5\%$)

Test de détection	TDO	TDO	TDK	TDK
Test du signe	TLO	TLK	TLO	TLK
C	0.121	0.121	0.302	0.302
$-l_1 \ l_2$	-2.94 2.94	-4.38 4.36	-2.94 2.94	-4.38 4.36
$ET_-^* \ ET_+^*$	18.8 4.71	18.9 4.70	18.8 4.71	18.9 4.70
p_1	70.8%	70.8%	75.4%	75.4%
p_2	98.4%	98.4%	98.8%	98.8%
p_3	35.0%	34.9%	40.1%	38.8%
p_4	100%	100%	100%	100%
$\bar{T}_- \ \bar{T}_+$	0.0 1.1	0.05 1.2	0.0 1.4	0.05 1.5

Tableau 3: résultats numériques pour l'exemple 6.2 ($\alpha_d = 5\%$, $\alpha_s = 5\%$)

Test de détection	TDO	TDO	TDK	TDK
Test du signe	TLO	TLK	TLO	TLK
C	0.144	0.144	0.144	0.144
$-l_1 \ l_2$	-2.94 2.94	-7.69 7.54	-2.94 2.94	-7.69 7.54
$ET_-^* \ ET_+^*$	2.35 0.117	2.40 0.115	2.35 0.117	2.40 0.115
p_1	85.5%	85.5%	87.0%	87.0%
p_2	95.3%	95.3%	92.5%	92.5%
p_3	66.8%	66.8%	69.3%	68.4%
p_4	95.8%	100.0%	90.0%	94.7%
$\bar{T}_- \ \bar{T}_+$	0.0 0.28	0.12 0.29	0.0 0.30	0.12 0.33

Tableau 4: résultats numériques pour l'exemple 6.3 ($\alpha_d = 5\%$, $\alpha_s = 5\%$)

Test de détection	TDO	TDO	TDK	TDK
Test du signe	TLO	TLK	TLO	TLK
C	0.165	0.165	0.165	0.165
$-l_1 \ l_2$	-4.70 2.37	-15.80 7.80	-4.70 2.37	-15.80 7.80
$ET_-^* \ ET_+^*$	0.44 0.13	2.46 0.44	0.44 0.13	2.46 0.44
p_1	86.6%	86.6%	88.2%	88.2%
p_2	93.5%	93.5%	87.5%	87.5%
p_3	71.2%	63.4%	76.1%	65.4%
p_4	82.4%	87.0%	71.1%	66.7%
$\bar{T}_- \ \bar{T}_+$	0.38 0.29	1.60 0.28	0.40 0.37	1.60 0.30

Tableau 5: résultats numériques pour l'exemple 6.4 ($\alpha_d = 5\%$, $\alpha_s = 5\%$)

A.2 Représentations graphiques

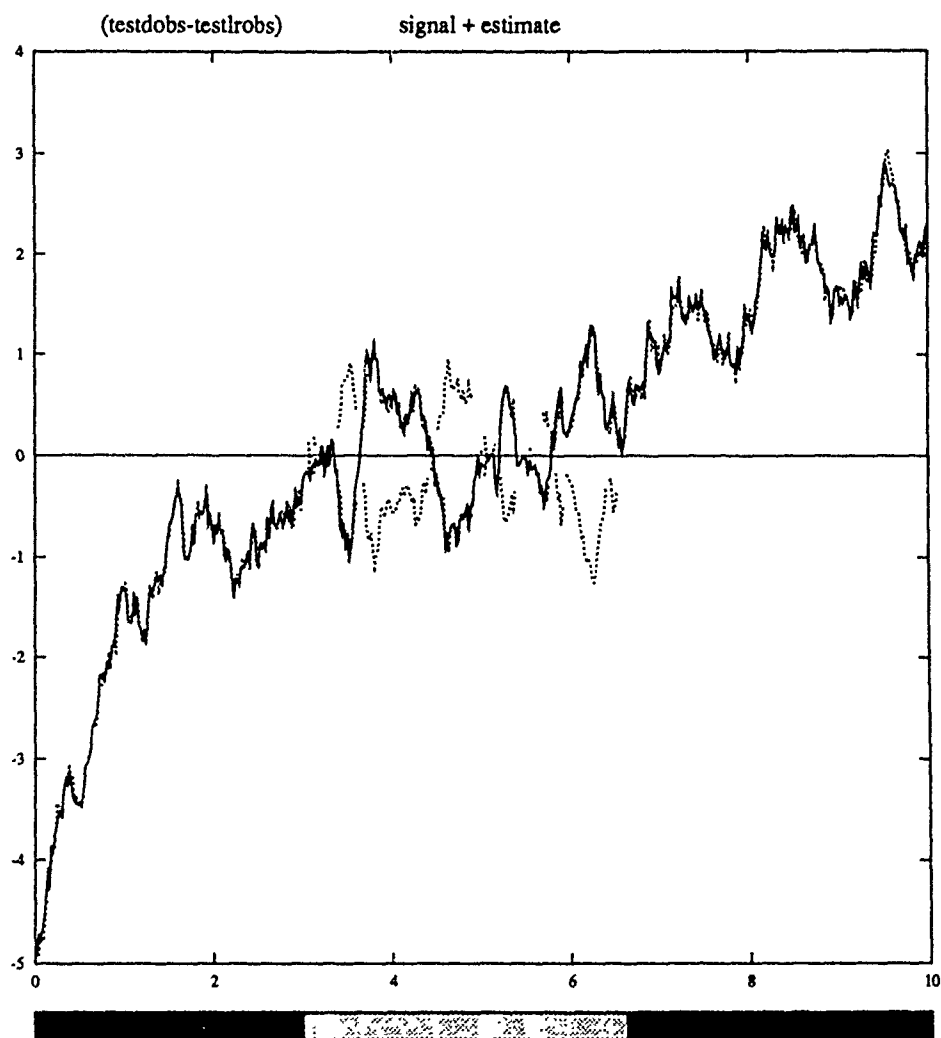


Figure 1: Sous ($HD2$), exemple 6.1: tests de détection des passages à zero et de décision du signe basés sur les observations.

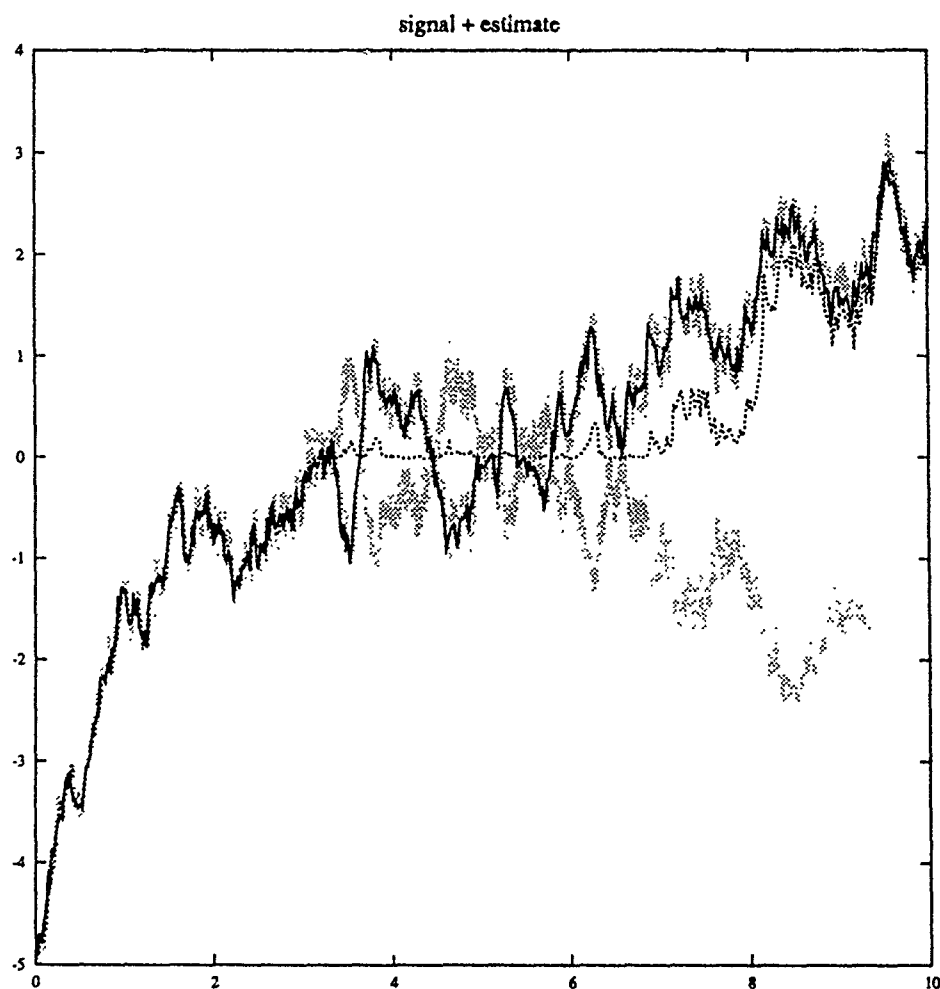


Figure 2: Sous ($HD2$), exemple 6.1: solution de l'équation de Zakai (le niveau de gris correspond à une région de confiance à 95%).

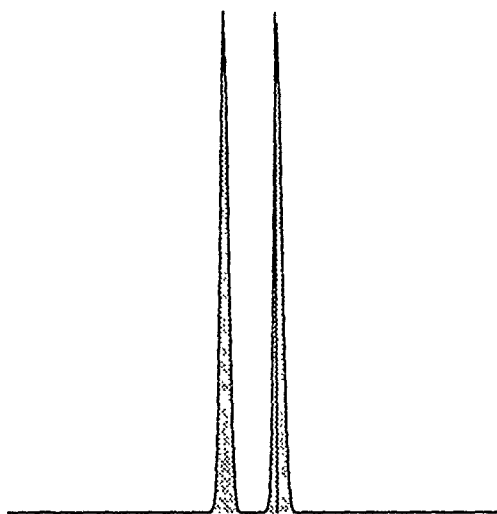


Figure 3: Sous ($HD2$), exemple 6.1: densité conditionnelle à l'instant $t = 6.5$

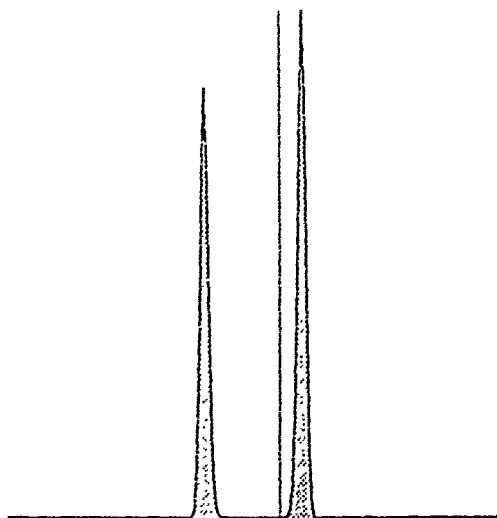


Figure 4: Sous ($HD2$), exemple 6.1: densité conditionnelle à l'instant $t = 7$

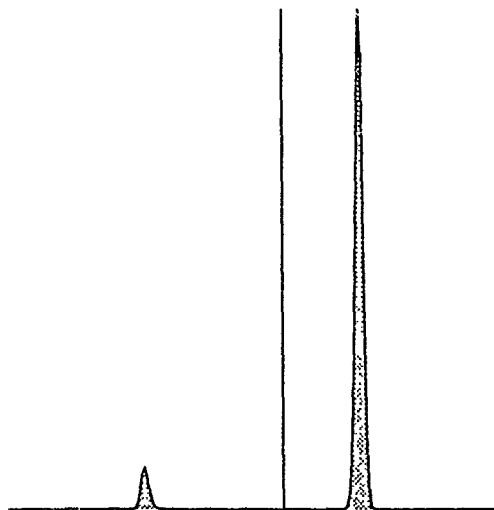


Figure 5: Sous ($HD2$), exemple 6.1: densité conditionnelle à l'instant $t = 8.48$

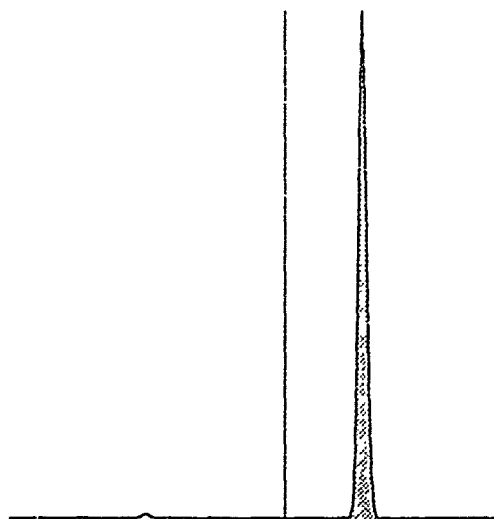


Figure 6: Sous ($HD2$), exemple 6.1: densité conditionnelle à l'instant $t = 10$

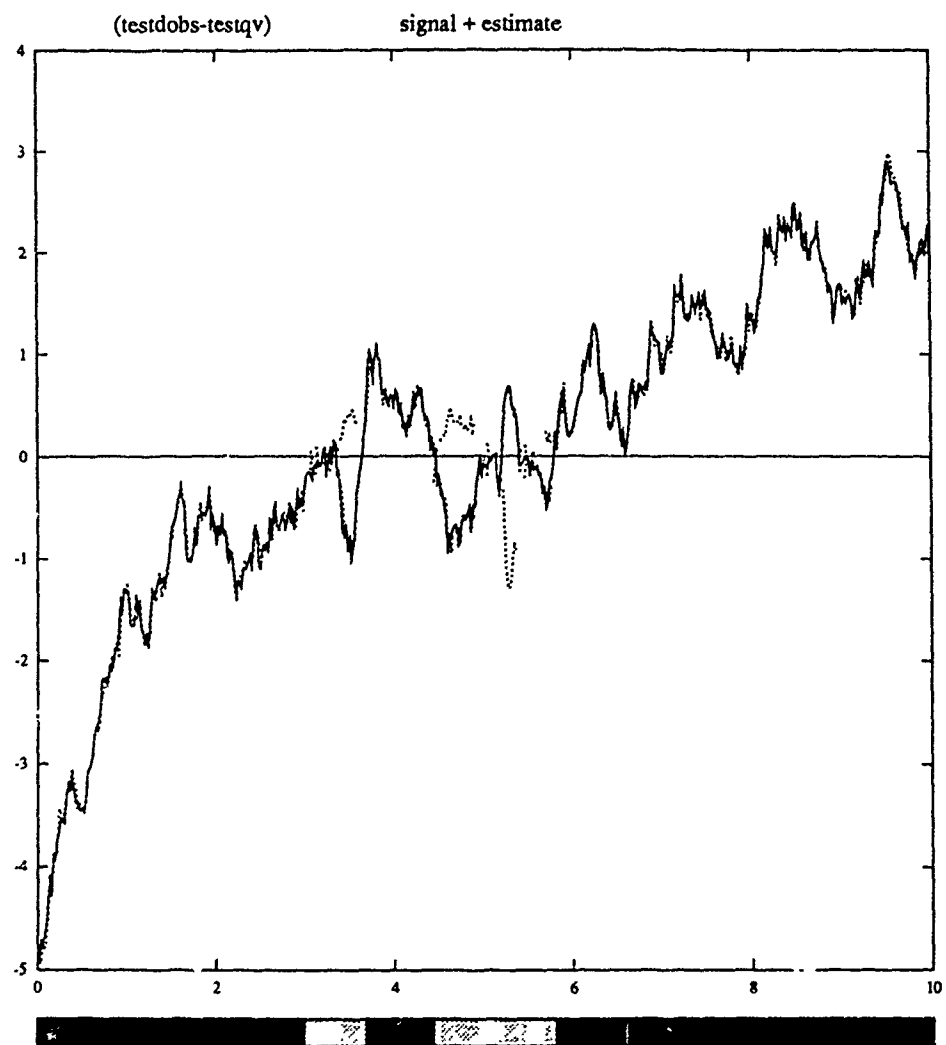


Figure 7: Sous ($HD1$), exemple 6.4: test de détection des passages à zero basé sur les observations et test de décision du signe basé sur la variation quadratique.

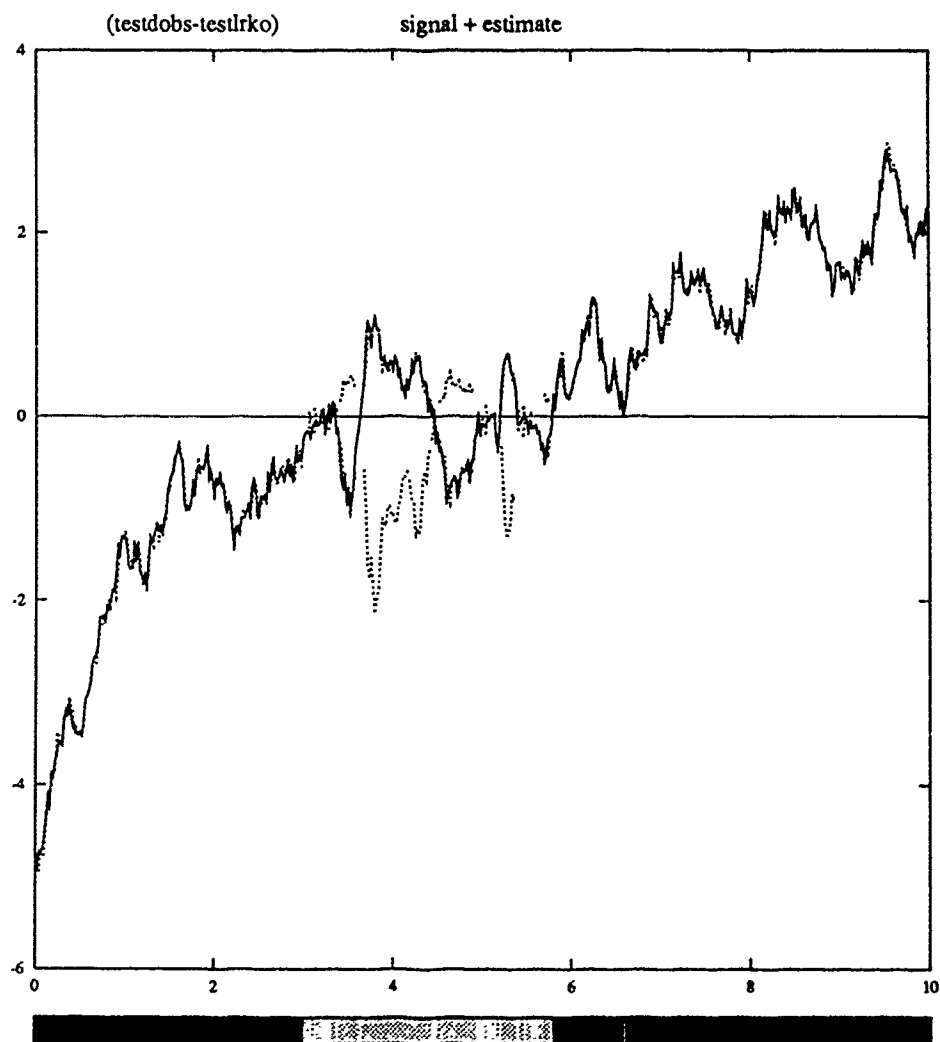


Figure 8: Sous ($HD1$), exemple 6.4: test de détection des passages à zero basé sur les observations et test de décision du signe basé sur la sortie des filtres de Kalman-Bucy.

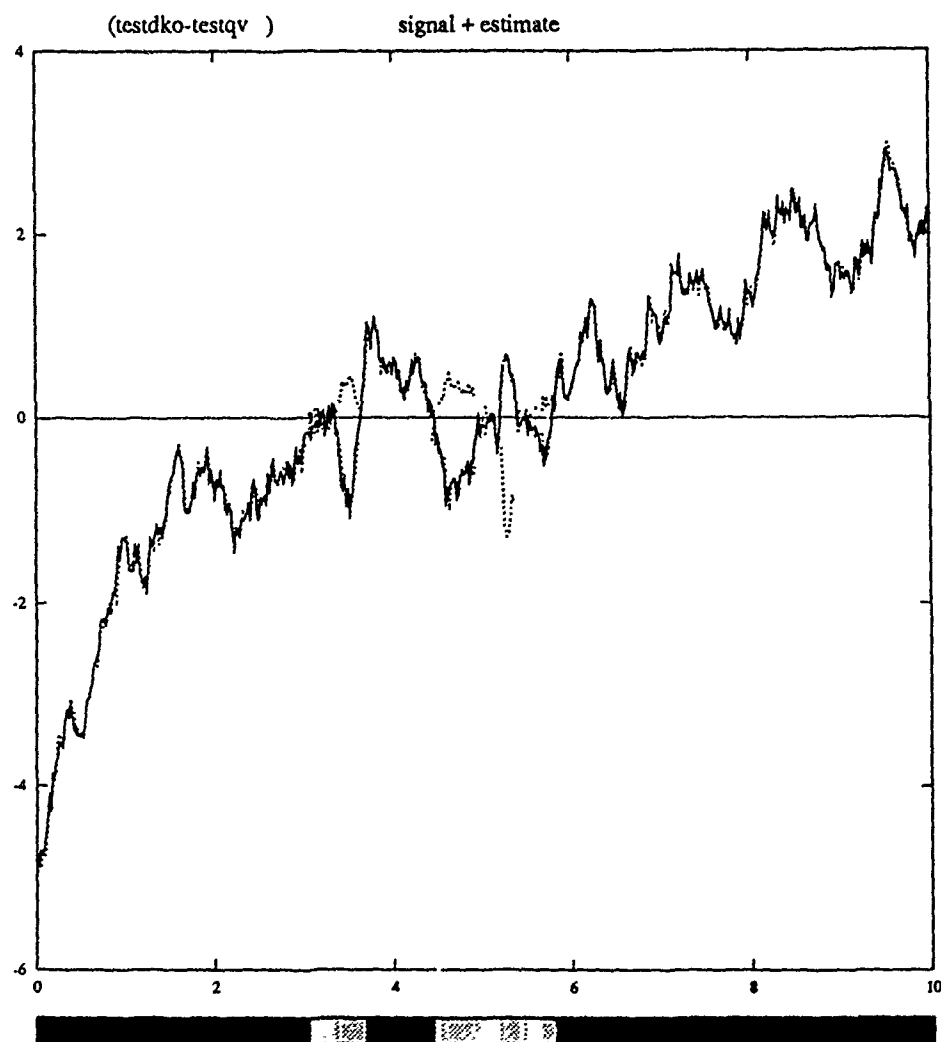


Figure 9: Sous ($HD1$), exemple 6.4: test de détection des passages à zero basé sur la sortie des filtres de Kalman-Bucy et test de décision du signe basé sur la variation quadratique.

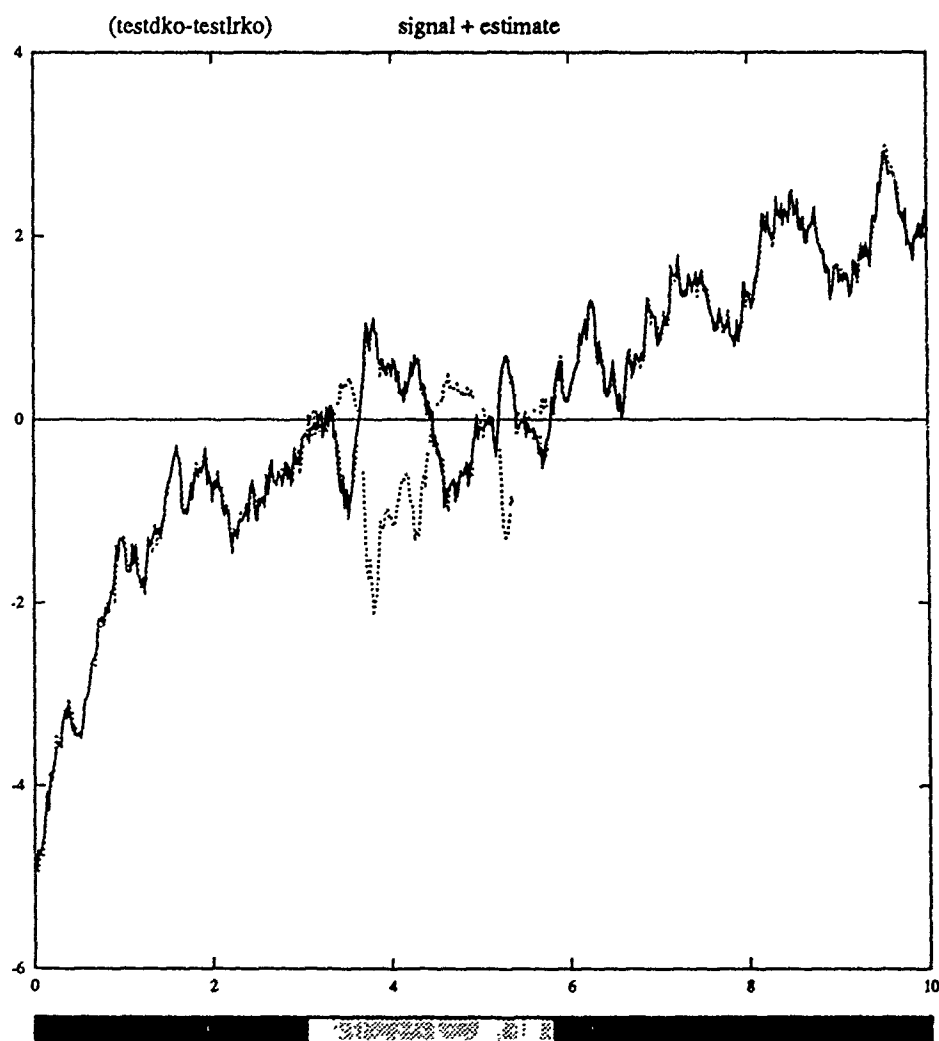


Figure 10: Sous (*HD1*), exemple 6.4: tests de détection des passages à zero et de décision du signe basés sur la sortie des filtres de Kalman-Bucy.

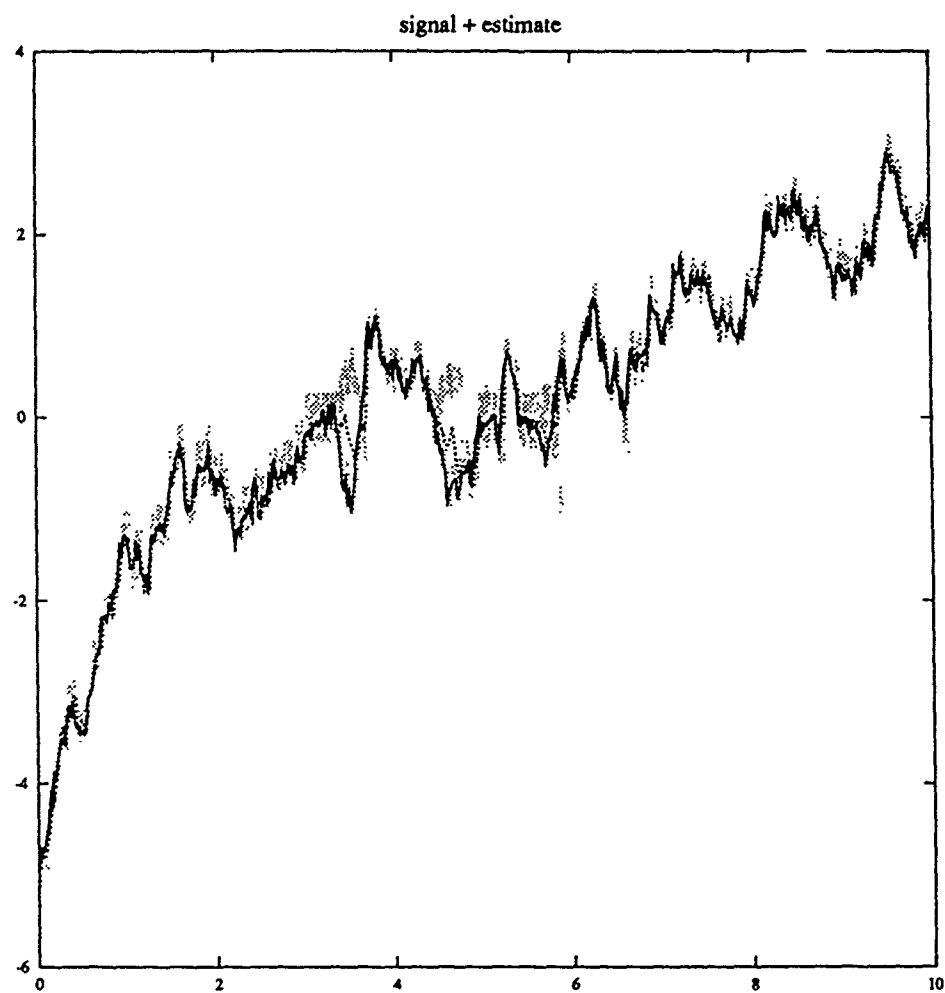


Figure 11: Sous ($HD1$), exemple 6.4: solution de l'équation de Zakai (le niveau de gris correspond à une région de confiance à 95%).

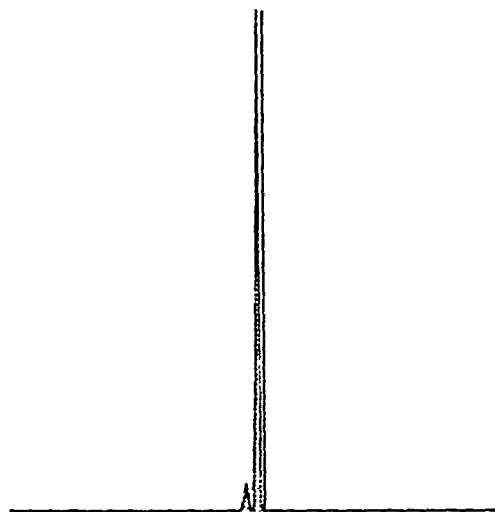


Figure 12: Sous ($HD1$), exemple 6.4: densité conditionnelle à l'instant $t = 6.60$

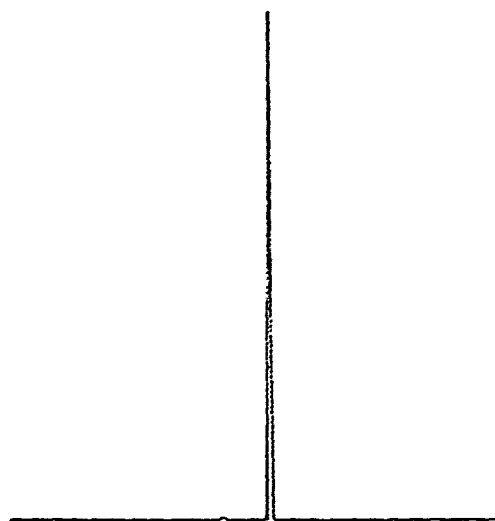


Figure 13: Sous ($HD1$), exemple 6.4: densité conditionnelle à l'instant $t = 6.86$

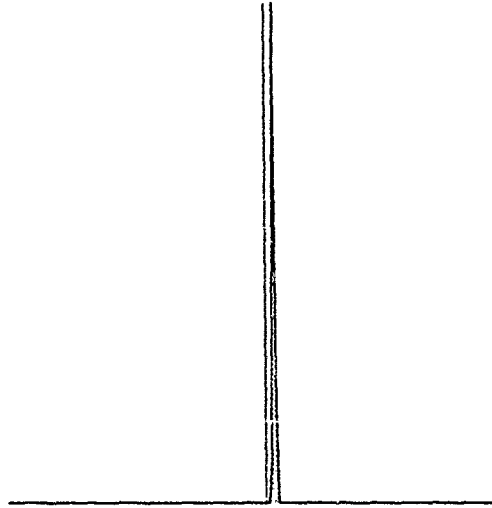


Figure 14: Sous ($HD1$), exemple 6.4: densité conditionnelle à l'instant $t = 6.88$

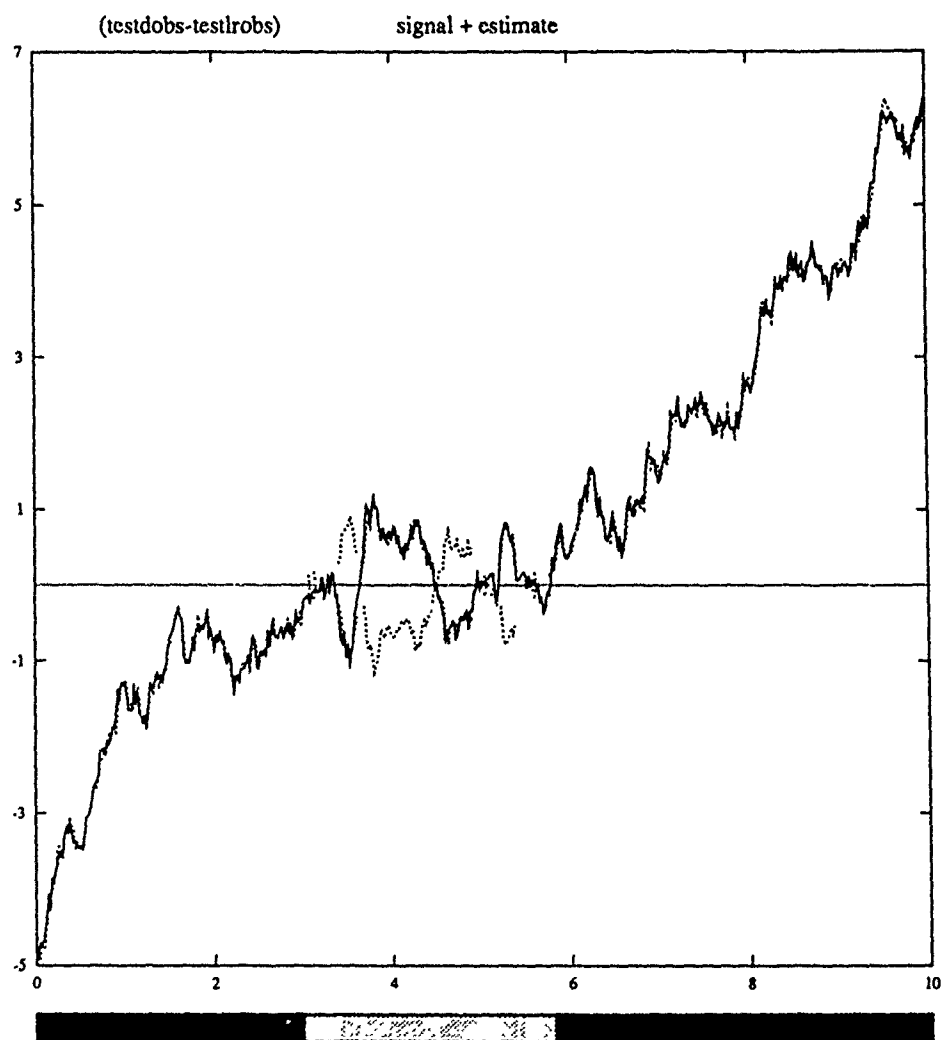


Figure 15: Exemple 6.6: test de détection des passages à zero basé sur les observations et test de décision du signe basé sur la variation quadratique.

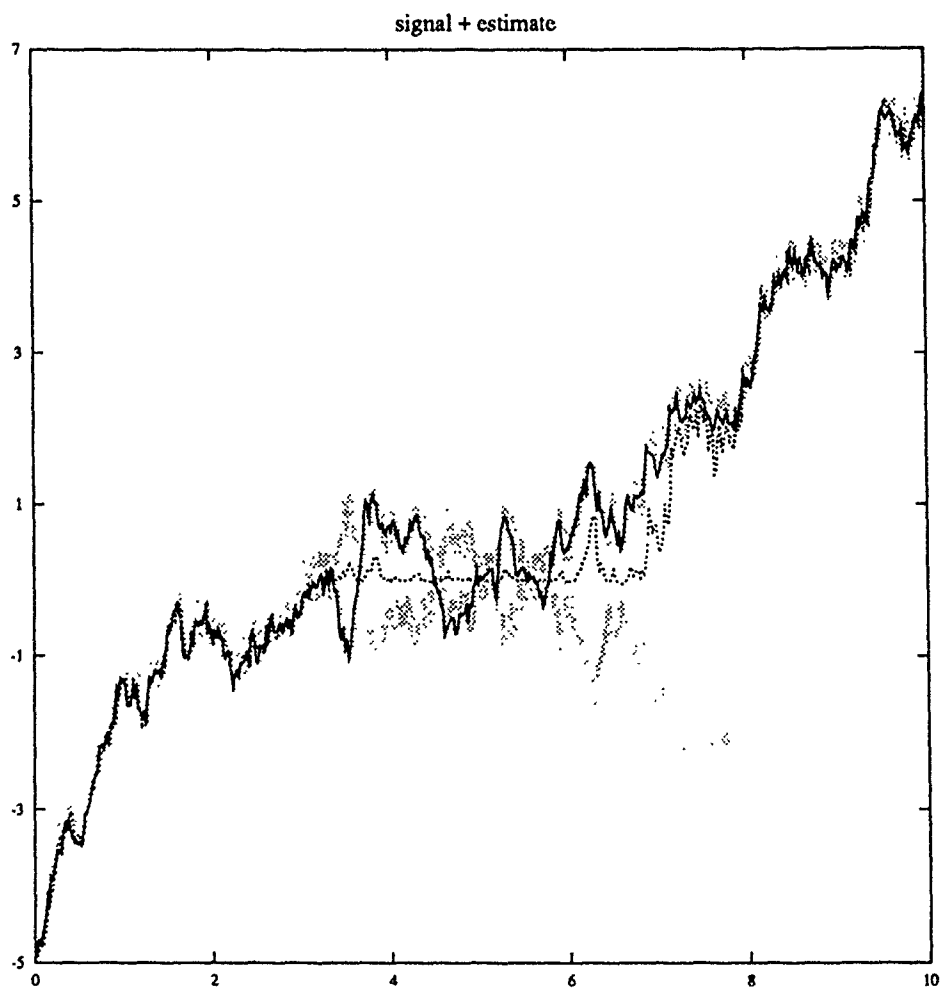


Figure 16: Exemple 6.6: solution de l'équation de Zakai (le niveau de gris correspond à une région de confiance à 95%).

3

Filtres approchés pour un
problème de filtrage non linéaire
d'une diffusion de dimension 2
mesurée par un processus de
dimension 1 faiblement bruité

FILTRES APPROCHÉS POUR UN PROBLÈME DE FILTRAGE
NON LINÉAIRE D'UNE DIFFUSION DE DIMENSION 2 MESURÉE
PAR UN PROCESSUS DE DIMENSION 1 FAIBLEMENT BRUITÉ

APPROXIMATE FILTERS FOR A NONLINEAR FILTERING
PROBLEM OF A TWO-DIMENSIONAL DIFFUSION MEASURED BY
A ONE-DIMENSIONAL PROCESS IN A LOW NOISE CHANNEL

Résumé

On fait une étude asymptotique d'un problème de filtrage non linéaire en temps continu, lorsque la variance du bruit d'observation tend vers 0. On considère un signal de dimension 2, dont une seule des composantes est bruitée, mesuré par un processus d'observation de dimension 1 faiblement bruité. Sous certaines hypothèses de détectabilité, on démontre que l'erreur de filtrage converge vers 0 et on donne une majoration pour cette vitesse de convergence. Les démonstrations font intervenir des résultats sur la stabilité des systèmes différentiels linéaires non autonomes.

Abstract

We study the asymptotic behaviour of a nonlinear continuous time filtering problem, as the variance of the observation noise tends to 0. We consider the case of a two-dimensional signal from which only one of the components is noisy, measured by a one-dimensional observation in a low noise channel. Under some detectability assumptions, we show that the filtering error converges to 0 and we give an upper bound for this converging rate. The proofs involve some results on the stability of linear non autonomous differential systems.

1 Introduction

Du à sa vaste application dans les domaines de l'ingénierie, le problème de filtrage d'un signal aléatoire à partir d'une observation bruitée a fait l'objet d'études de la part de plusieurs auteurs. En particulier, le cas d'une observation faiblement bruitée a été abondamment traité, pendant ces dernières années. On peut trouver, dans la littérature, des travaux consacrés à la recherche de filtres approchés asymptotiquement efficaces quand le bruit tend vers 0. Parmi ces travaux on remarque un premier groupe où le cas d'un système unidimensionnel ayant une fonction d'observation injective est traité (cf. [KBS], [Picard-86], [Bensoussan]). L'étude du cas multidimensionnel paraît plus tard avec [Picard-88]. Dans le cas d'une fonction d'observation non injective, des efforts ont été faits pour résoudre certaines classes de problèmes. D'abord avec [Fleming-Pardoux], qui ont traité le cas unidimensionnel avec une fonction d'observation monotone par morceaux et plus récemment [YBS] et [Picard-89]. Dans [YBS], les auteurs étudient formellement le filtrage d'un signal de dimension 2 mesuré par un processus d'observation unidimensionnel. Dans [Picard-89], l'auteur considère le problème général du filtrage d'un signal multidimensionnel au vu des réalisations d'un processus d'observation multidimensionnel et il énonce et démontre des résultats sur l'efficacité du filtre de Kalman étendu quand le bruit est faible.

Dans ce travail, on considère un système pour lequel la dynamique est modélisée par un processus de diffusion de dimension 2, dont seulement une des composantes est bruitée, et l'observation est modélisée par un processus de dimension 1, faiblement bruité, mesurant l'autre composante.

On considère un espace de probabilité $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$.

On s'intéresse au problème suivant. On veut estimer, à l'instant t , un signal $\{X_t\}$ de dimension 2 donné par l'équation d'Itô

$$\begin{cases} dx_1(t) = f_1(x_1(t), x_2(t)) dt \\ dx_2(t) = f_2(x_1(t), x_2(t)) dt + \sigma(x_1(t), x_2(t)) dw(t), \end{cases} \quad (1)$$

avec condition initiale $X_0 = (x_1(0), x_2(0))$, au vu des observations $\sigma\{y_s, 0 \leq s \leq t\}$ données par l'équation

$$dy(t) = h(x_1(t)) dt + \varepsilon d\bar{w}(t), \quad (2)$$

où $\{w\}$ et $\{\bar{w}\}$ sont des processus de Wiener standards indépendants à valeurs dans \mathcal{R} et ε est un paramètre supposé "petit", $\varepsilon > 0$.

On cherche à démontrer que, sous certaines hypothèses de régularité sur les fonctions $f = (f_1, f_2)$, σ et h , ce signal peut être estimé par un filtre de dimension finie.

Dans [YBS], les auteurs considèrent ce problème, entre autres. Ils obtiennent, dans une situation stationnaire, des développements asymptotiques pour le filtre optimal et l'erreur quadratique moyenne. De cette étude, ils déduisent un filtre sous-optimal de dimension finie asymptotiquement performant.

Nous allons démontrer que l'erreur de filtrage est asymptotiquement nulle quand $\varepsilon \rightarrow 0$ et qu'il existe un filtre approché de dimension finie pour ce problème. Notre point de départ sera le filtre approché proposé par [YBS]. Nous obtiendrons une estimation de la vitesse de convergence confirmant le résultat de ces auteurs. La méthode de démonstration utilisée est inspirée par le contenu de [Picard-89].

Ce travail est organisé comme suit. Dans une première étape on étudie le cas où la fonction f_1 est linéaire par rapport à x_2 et la fonction h est aussi linéaire. Ceci est fait dans la section 2. Dans la section 3, on étudie le cas plus général d'un système modélisé par les équations (1) et (2). Pour cela on se ramène, par une transformation de coordonnées, au cas traité dans la section précédente. On termine notre étude avec quelques commentaires sur le problème.

On utilisera, par la suite, la définition suivante:

Définition 1.1 Soit A_t une famille (indexée par ε) de processus mesurables bornés à valeurs dans l'espace des matrices 2×2 . Soit ζ_t la solution (à valeurs matricielles) de l'équation différentielle

$$\begin{cases} \dot{\zeta}_t = A_t \zeta_t \\ \zeta_0 = I. \end{cases} \quad (3)$$

Soit $\zeta_{ts} = \zeta_t \zeta_s^{-1}$. On dit que A_t est exponentiellement stable avec taux d'ordre $1/\sqrt{\varepsilon}$ s'il existe des constantes $c, C > 0$ telles que

$$\forall s \leq t, \|\zeta_{ts}\| \leq C \exp\left\{-\frac{c(t-s)}{\sqrt{\varepsilon}}\right\},$$

où $\|\cdot\|$ représente la norme $\|A\| = \sup_{|u|=1} |Au|$ et $|\cdot|$ est la norme euclidienne dans \mathcal{R}^2 .

Notation 1.2 On notera

$$R_t = \begin{pmatrix} r_1(t) \\ r_2(t) \end{pmatrix}$$

un processus à valeurs dans \mathcal{R}^2 .

Les constantes strictement positives indépendantes de ε intervenant dans les calculs seront notées c ou C , sans expliciter leur valeur.

Si A est une matrice on notera A^ sa transposée.*

Par abus de notation, on supprimera quelques fois la dépendance en temps dans les objets mathématiques ayant un indice inférieur.

2 Le cas particulier d'une dérive linéaire par rapport à x_2 et d'une fonction d'observation linéaire

On considère le cas particulier du système

$$\begin{cases} dx_1(t) = f_1(x_1(t), x_2(t)) dt \\ dx_2(t) = f_2(x_1(t), x_2(t)) dt + \sigma(x_1(t), x_2(t)) dw(t) \\ dy(t) = Hx_1(t) dt + \varepsilon d\bar{w}(t), \end{cases} \quad (4)$$

ayant comme hypothèses

(H1) $H > 0$ et σ est une fonction bornée

(H2) f_1 et f_2 sont des fonctions C^1 à dérivées bornées et

$$\frac{\partial f_1}{\partial x_2} = F_{12} > 0.$$

On considère le filtre $\{M_t\}$ donné par l'équation d'Itô

$$\begin{cases} dm_1(t) = f_1(m_1(t), m_2(t)) dt + \sqrt{\frac{2\tilde{\sigma}F_{12}}{\varepsilon H}} [dy(t) - Hm_1(t) dt] \\ dm_2(t) = f_2(m_1(t), m_2(t)) dt + \frac{\tilde{\sigma}}{\varepsilon} [dy(t) - Hm_1(t) dt], \end{cases} \quad (5)$$

avec condition initiale M_0 et où $\tilde{\sigma}$ est une constante strictement positive.

Pour $\sigma = c^{1\varepsilon} > 0$, le filtre proposé dans [YBS] coïncide avec ce filtre quand on prend $\tilde{\sigma} = \sigma$.

On cherche à estimer l'erreur de ce filtre, i.e., pour $p \geq 1$, on s'intéresse à des majorations de $E[\|X_t - M_t\|^{2p}]$. Soit $\{Z_t\}$ le processus

$$Z_t = \begin{pmatrix} 1/\sqrt{\varepsilon} & 0 \\ 0 & 1 \end{pmatrix} (X_t - M_t).$$

On commence par écrire ce processus comme solution d'une équation différentielle, dont la dérive sera définie par une certaine matrice A . De l'étude sur la stabilité de cette matrice on déduit, en utilisant quelques propriétés élémentaires du calcul matriciel, l'estimation de l'erreur du filtre.

2.1 Détermination d'une équation différentielle stochastique pour le processus $\{Z_t\}$

Des équations (4) et (5), on obtient l'équation différentielle stochastique

$$dZ_t = \begin{pmatrix} 1/\sqrt{\varepsilon} & 0 \\ 0 & 1 \end{pmatrix} [f(X_t) - f(M_t)] dt - H(x_1 - m_1) \begin{pmatrix} 1/\sqrt{\varepsilon} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{\frac{2\tilde{\sigma}F_{12}}{\varepsilon H}} \\ \tilde{\sigma}/\varepsilon \end{pmatrix} dt \\ + \begin{pmatrix} 1/\sqrt{\varepsilon} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -\sqrt{\frac{2\tilde{\sigma}F_{12}\varepsilon}{H}} d\bar{w}_t \\ \sigma(X_t) dw_t - \tilde{\sigma} d\bar{w}_t \end{pmatrix}.$$

On applique la formule de Taylor à la fonction f :

$$f(X_t) - f(M_t) = F(\xi_t, \mu_t)(X_t - M_t),$$

où $\{\xi_t\}$ et $\{\mu_t\}$ sont des processus à valeurs dans \mathcal{R}^2 qui dépendent de $\{X_t\}$ et de $\{M_t\}$ et

$$F(\xi_t, \mu_t) \triangleq \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\xi_t) & \frac{\partial f_1}{\partial x_2}(\xi_t) \\ \frac{\partial f_2}{\partial x_1}(\mu_t) & \frac{\partial f_2}{\partial x_2}(\mu_t) \end{pmatrix}.$$

On obtient alors que le processus $\{Z_t\}$ est solution de l'équation différentielle stochastique

$$dZ_t = A_t Z_t dt + \begin{pmatrix} -\sqrt{\frac{2\tilde{\sigma}F_{12}}{H}} d\bar{w}_t \\ \sigma(X_t) dw_t - \tilde{\sigma} d\bar{w}_t \end{pmatrix}, \quad (6)$$

où

$$A_t \triangleq \begin{pmatrix} F_{11}(\xi_t, \mu_t) - \sqrt{\frac{2\tilde{\sigma}F_{12}H}{\varepsilon}} & F_{12}/\sqrt{\varepsilon} \\ F_{21}(\xi_t, \mu_t)\sqrt{\varepsilon} - H\tilde{\sigma}/\sqrt{\varepsilon} & F_{22}(\xi_t, \mu_t) \end{pmatrix} \quad (7)$$

est un processus à valeurs dans l'espace des matrices 2×2 .

2.2 Etude de la stabilité du processus matriciel $\{A_t\}$

L'étude de la stabilité des systèmes linéaires non autonomes sera fondamentale dans l'obtention des estimations pour l'erreur de filtrage.

On établit d'abord le lemme suivant:

Lemme 2.1 Soit

$$B = \begin{pmatrix} 2\tilde{\sigma}H & -\sqrt{2\tilde{\sigma}F_{12}H} \\ -\sqrt{2\tilde{\sigma}F_{12}H} & 2F_{12} \end{pmatrix}$$

une matrice symétrique et

$$\lambda = \frac{1}{2} \sqrt{\frac{2\tilde{\sigma}F_{12}H}{\varepsilon}}.$$

Alors B est définie positive et

$$BA_t + A_t^*B \leq -\lambda B, \forall t.$$

Preuve

Le fait que B est une matrice définie positive est évident. Il suffit de remarquer que $\text{trace}B > 0$ et $\det B = 2\tilde{\sigma}F_{12}H > 0$.

Soit $D_t = BA_t + A_t^*B + \lambda B$. On veut démontrer que D_t est semi-définie négative, pour tout t .

Un calcul élémentaire matriciel nous donne les expressions de la trace et du déterminant de D_t :

$$\begin{aligned} \text{trace}D_t &= 4\tilde{\sigma}H a_{11}(t) - 2\sqrt{2\tilde{\sigma}F_{12}H} a_{21}(t) + \tilde{\sigma}H \sqrt{\frac{2\tilde{\sigma}F_{12}H}{\varepsilon}} \\ &\quad - 2\sqrt{2\tilde{\sigma}F_{12}H} a_{12}(t) + 4F_{12} a_{22}(t) + \sqrt{\frac{2\tilde{\sigma}F_{12}H}{\varepsilon}} F_{12} \\ \det D_t &= \left[4\tilde{\sigma}H a_{11}(t) - 2\sqrt{2\tilde{\sigma}F_{12}H} a_{21}(t) + \tilde{\sigma}H \sqrt{\frac{2\tilde{\sigma}F_{12}H}{\varepsilon}} \right] \\ &\quad \cdot \left[-2\sqrt{2\tilde{\sigma}F_{12}H} a_{12}(t) + 4F_{12} a_{22}(t) + \sqrt{\frac{2\tilde{\sigma}F_{12}H}{\varepsilon}} F_{12} \right] \\ &\quad - \left[-\sqrt{2\tilde{\sigma}F_{12}H} a_{11}(t) + 2F_{12} a_{21}(t) + 2\tilde{\sigma}H a_{12}(t) - \sqrt{2\tilde{\sigma}F_{12}H} a_{22}(t) \right. \\ &\quad \left. - \frac{\tilde{\sigma}F_{12}H}{\sqrt{\varepsilon}} \right]^2. \end{aligned}$$

Etant donné l'ordre de grandeur des éléments de la matrice A_t et, en utilisant les hypothèses (H1) et (H2), on obtient

$$\begin{aligned} \text{trace} D_t &= -(\tilde{\sigma} H + F_{12}) \sqrt{\frac{2 \tilde{\sigma} F_{12} H}{\varepsilon}} + \mathcal{O}(1) \leq 0, \forall t \\ \det D_t &= \frac{\tilde{\sigma}^2 H^2 F_{12}^2}{\varepsilon} + \mathcal{O}\left(\frac{1}{\sqrt{\varepsilon}}\right) > 0, \forall t. \end{aligned}$$

On peut maintenant démontrer la

Proposition 2.2 *Le processus A_t est exponentiellement stable avec taux d'ordre $1/\sqrt{\varepsilon}$.*

Preuve

Soient B et λ définis par le lemme 2.1. Soit L la forme quadratique associée à la matrice B :

$$L(X) = X^* B X.$$

Soit ζ la solution de l'équation (3). On note ζ_{ts}^i la i^{me} colonne de ζ_t . On a donc que ζ_{ts}^i est solution de l'équation

$$\dot{\zeta}_{ts}^i = A_t \zeta_{ts}^i.$$

Alors $L_t = L(\zeta_{ts}^i)$ est une fonction de Lyapunov, puisque

- i) $L(0) = 0$, $L(X) > 0$ si $X \neq 0$
- ii) $\frac{dL}{dt}(\zeta_{ts}^i) = \zeta_{ts}^{i*} (B A_t + A_t^* B) \zeta_{ts}^i$.

D'après le lemme 2.1, on a la majoration

$$\frac{dL}{dt}(\zeta_{ts}^i) \leq -\lambda \zeta_{ts}^{i*} B \zeta_{ts}^i = -\lambda L(\zeta_{ts}^i).$$

Vu que B est une matrice bornée, on a

$$L(\zeta_{ts}^i) \leq C e^{-\lambda(t-s)}$$

et, puisque B est uniformément elliptique ($B \geq cI$), on a

$$\|\zeta_{ts}^i\| \leq C e^{-\lambda(t-s)},$$

ce qui est l'estimation recherchée.

2.3 Estimation de l'erreur de filtrage

On établit le résultat suivant:

Théorème 2.3 Soit p un entier, $p \geq 1$. Si $E[\|X_0 - M_0\|^{2p}] \leq C < +\infty$ alors on a les estimations

$$\begin{aligned} E[(x_1(t) - m_1(t))^{2p}] &\leq C_p e^{-c_p t/\sqrt{\varepsilon}} + C_p \varepsilon^{3p/2} \\ E[(x_2(t) - m_2(t))^{2p}] &\leq \frac{C_p}{\varepsilon^p} e^{-c_p t/\sqrt{\varepsilon}} + C_p \varepsilon^{p/2}. \end{aligned}$$

La preuve de ce théorème nécessite 2 lemmes qu'on énonce et démontre.

On définit $\{Q_t\}$ comme étant le processus à valeurs dans l'espace des matrices (symétriques définies positives) solution de l'équation

$$\begin{cases} \dot{Q}_t = A_t Q_t + Q_t A_t^* + I \\ Q_0 = q_0 \sqrt{\varepsilon} I. \end{cases} \quad (8)$$

Lemme 2.4 La matrice $\{Q_t\}$ est bornée, d'ordre $\mathcal{O}(\sqrt{\varepsilon})$.

Preuve

Puisque Q_t est la solution de l'équation (8), on connaît sa forme explicite:

$$Q_t = \zeta_t Q_0 \zeta_t^* + \int_0^t \zeta_{ts} \zeta_{ts}^* ds$$

où $\{\zeta_t\}$ est la solution de l'équation (3).

D'après la proposition 2.2,

$$\forall s \leq t, \|\zeta_{ts}\| \leq C e^{-c(t-s)/\sqrt{\varepsilon}}.$$

De la majoration de la norme de l'intégrale on obtient la majoration suivante pour la norme de Q_t :

$$\|Q_t\| \leq C \sqrt{\varepsilon}.$$

■

On considère le processus $V_t = Q_t^{-1}$.

Ce processus est solution de l'équation de Riccati

$$\begin{cases} \dot{V}_t = -V_t A_t - A_t^* V_t - V_t^2 \\ V_0 = Q_0^{-1}. \end{cases} \quad (9)$$

Lemme 2.5 *La matrice V_t est bornée et elliptique uniformément par rapport à t . Plus précisément,*

$$V_t \geq \frac{c}{\sqrt{\varepsilon}} I \text{ et } \text{trace} V_t \leq \frac{C}{\sqrt{\varepsilon}}.$$

Preuve

Du lemme 2.4 vient immédiatement que

$$V_t \geq \frac{1}{\|Q_t\|} I \geq \frac{c}{\sqrt{\varepsilon}} I.$$

De l'équation (9) on obtient, en utilisant les propriétés

$$\begin{aligned} |\text{trace}(V_t A_t)| &\leq \|A_t\| \text{trace} V_t \\ \text{trace}(V_t^2) &\geq \frac{1}{2} (\text{trace} V_t)^2 \text{ (cf. [Saniuk-Rhodes])} \end{aligned}$$

et le fait que $\|A_t\| \leq C/\sqrt{\varepsilon}$, une inéquation différentielle pour la trace de V_t :

$$\frac{d}{dt}(\text{trace} V_t) \leq \frac{C}{\sqrt{\varepsilon}} (\text{trace} V_t) - \frac{1}{2} (\text{trace} V_t)^2.$$

La résolution de cette inéquation permet d'obtenir la majoration:

$$\text{trace} V_t \leq \frac{1}{c\sqrt{\varepsilon} + (1/(\text{trace} V_0) - c\sqrt{\varepsilon}) e^{-Ct/\sqrt{\varepsilon}}},$$

d'où le lemme. ■

On est prêt à démontrer le théorème 2.3.

Preuve (du théorème 2.3)

Soit

$$U_t \triangleq \begin{pmatrix} 0 & -\sqrt{\frac{2\tilde{\sigma}F_{12}}{H}} \\ \sigma(X_t) & -\tilde{\sigma} \end{pmatrix}.$$

D'après la formule d'Itô,

$$d(Z_t^* V_t Z_t) = -Z_t^* V_t^2 Z_t dt + \text{trace}(U_t^* V_t U_t) dt + 2 Z_t^* V_t U_t \begin{pmatrix} dw_t \\ d\tilde{w}_t \end{pmatrix}.$$

Soit $p \geq 1$. D'après le lemme de Gronwall, les moments d'ordre p de $Z_t^* V_t Z_t$ sont finis. On applique la formule d'Itô à $(Z_t^* V_t Z_t)^p$:

$$\begin{aligned} \frac{d}{dt} (Z_t^* V_t Z_t)^p &= -p (Z_t^* V_t Z_t)^{p-1} Z_t^* V_t^2 Z_t dt + p (Z_t^* V_t Z_t)^{p-1} \text{trace}(U_t^* V_t U_t) dt \\ &\quad + 2p(p-1) (Z_t^* V_t Z_t)^{p-1} |\nu_t|^2 dt + 2p (Z_t^* V_t Z_t)^{p-1} (Z_t^* V_t U_t) \left(\frac{dw_t}{d\bar{w}_t} \right), \end{aligned}$$

où

$$\nu_t = \frac{1}{(Z_t^* V_t Z_t)^{1/2}} U_t^* V_t Z_t.$$

On obtient alors l'expression

$$\begin{aligned} \frac{d}{dt} \mathbf{E}[(Z_t^* V_t Z_t)^p] &= -p \mathbf{E}[(Z_t^* V_t Z_t)^{p-1} Z_t^* V_t^2 Z_t] + p \mathbf{E}[(Z_t^* V_t Z_t)^{p-1} \text{trace}(U_t^* V_t U_t)] \\ &\quad + 2p(p-1) \mathbf{E}[(Z_t^* V_t Z_t)^{p-1} |\nu_t|^2]. \end{aligned}$$

D'après le lemme 2.5, on a

$$Z_t^* V_t^2 Z_t \geq \alpha_\varepsilon Z_t^* V_t Z_t$$

avec $\alpha_\varepsilon = c/\sqrt{\varepsilon}$ et l'inégalité de Cauchy-Schwartz entraine

$$|\nu_t|^2 \leq \text{trace}(U_t^* V_t U_t).$$

On obtient l'inégalité

$$\frac{d}{dt} \mathbf{E}[(Z_t^* V_t Z_t)^p] \leq -p \alpha_\varepsilon \mathbf{E}[(Z_t^* V_t Z_t)^p] + p(2p-1) \mathbf{E}[(Z_t^* V_t Z_t)^{p-1} \text{trace}(U_t^* V_t U_t)].$$

Puisque $\|U_t\|$ est bornée d'ordre $\mathcal{O}(1)$, on a

$$\text{trace}(U_t^* V_t U_t) \leq C_0 \alpha_\varepsilon.$$

On utilise la propriété

$$\forall \text{ suite } c_p > 0 \exists \text{ suite } C_p : x^{p-1}y \leq c_p x^p + C_p y^p, \forall x, y \geq 0.$$

On prend $c_p = (1-c)/[(2p-1)C_0]$, où c est une constante, $c < 1$.

On obtient l'inégalité

$$\frac{d}{dt} \mathbf{E}[(Z_t^* V_t Z_t)^p] \leq -\frac{c}{\sqrt{\varepsilon}} p \mathbf{E}[(Z_t^* V_t Z_t)^p] + \frac{C'_p}{\sqrt{\varepsilon}} \mathbf{E} \left[\left(\frac{\text{trace}(U_t^* V_t U_t)}{C_0 \alpha_\varepsilon} \right)^p \right],$$

où, on le rappelle, $E \left[\left(\frac{\text{trace}(U_t^* V_t U_t)}{C_0 \alpha_\varepsilon} \right)^p \right]$ est borné.

En résolvant cette inéquation différentielle on obtient la majoration

$$E[(Z_t^* V_t Z_t)^p] \leq C_p + E[(Z_0^* V_0 Z_0)^p] e^{-c p t / \sqrt{\varepsilon}},$$

où, vu que $E[\|X_0 - M_0\|^{2p}]$ est supposé fini, $E[(Z_0^* V_0 Z_0)^p] \leq C/\varepsilon^{3p/2}$.

Du lemme 2.5, en utilisant le fait que

$$\begin{aligned} \frac{1}{\varepsilon^p} E[(x_1(t) - m_1(t))^{2p}] + E[(x_2(t) - m_2(t))^{2p}] &\leq E[(Z_t^* Z_t)^p] \\ &\leq C_p \varepsilon^{p/2} + \frac{C}{\varepsilon^p} e^{-c p t / \sqrt{\varepsilon}}, \end{aligned}$$

on déduit le théorème. ■

Remarque 2.6 En plus de l'ordre de grandeur de l'erreur de filtrage on a obtenu que l'effet de la condition initiale disparaît à une vitesse exponentielle.

3 Le cas général

On considère le système

$$\begin{cases} dx_1(t) = f_1(x_1(t), x_2(t)) dt \\ dx_2(t) = f_2(x_1(t), x_2(t)) dt + \sigma(x_1(t), x_2(t)) dw(t) \\ dy(t) = h(x_1(t)) dt + \varepsilon d\bar{w}(t), \end{cases} \quad (10)$$

ayant comme hypothèses

(H3) h est une fonction C^2 , h' est bornée strictement positive, h'' est bornée et h^{-1} est Lipschitzienne

(H4) f est une fonction C^2 , f_1 est inversible par rapport à x_2 , d'inverse Lipschitzienne.

On note \bar{g} cette inverse

$$x_2 = \bar{g}(x_1, \bar{x}_2) \leftrightarrow \bar{x}_2 = f_1(x_1, x_2)$$

et g la fonction définie par

$$g(\bar{x}_1, \bar{x}_2) = \bar{g}(h^{-1}(\bar{x}_1), \bar{x}_2/h'(h^{-1}(\bar{x}_1)))$$

(H5) $h'f_1$ est une fonction Lipschitzienne

(H6) la fonction

$$\sigma(x_1, x_2) h'(x_1) \frac{\partial f_1}{\partial x_2}(x_1, x_2)$$

est bornée

(H7) la fonction

$$\begin{aligned} & h''(x_1) f_1^2(x_1, x_2) + h'(x_1) \frac{\partial f_1}{\partial x_1}(x_1, x_2) f_1(x_1, x_2) \\ & + h'(x_1) \frac{\partial f_1}{\partial x_2}(x_1, x_2) f_2(x_1, x_2) + \frac{\sigma^2(x_1, x_2)}{2} h'(x_1) \frac{\partial^2 f_1}{\partial x_2^2}(x_1, x_2) \end{aligned}$$

est différentiable et ses dérivées partielles sont bornées.

On va démontrer que, pour ce système, le théorème 2.3 énoncé dans la section précédente est encore vrai pour $p = 1$. En effet, par une transformation de coordonnées, on peut se ramener à un système du type traité dans la section 2.

Des estimations de l'erreur pour ce système transformé, on va pouvoir déduire des estimations analogues pour le système initial.

On considère la transformation

$$\begin{cases} \bar{x}_1(t) = h(x_1(t)) \\ \bar{x}_2(t) = h'(x_1(t)) f_1(x_1(t), x_2(t)). \end{cases}$$

Le nouveau système est, d'après la formule d'Itô,

$$\begin{cases} d\bar{x}_1(t) = \bar{x}_2(t) dt \\ d\bar{x}_2(t) = \bar{f}(\bar{x}_1(t), \bar{x}_2(t)) dt + \bar{\sigma}(\bar{x}_1(t), \bar{x}_2(t)) dw(t) \\ dy(t) = \bar{x}_1(t) dt + \varepsilon d\bar{w}(t), \end{cases} \quad (11)$$

où

$$\begin{aligned} \bar{f}(\bar{x}_1, \bar{x}_2) &= h''(h^{-1}(\bar{x}_1)) f_1^2(h^{-1}(\bar{x}_1), \bar{g}(\bar{x}_1, \bar{x}_2)) \\ &\quad + h'(h^{-1}(\bar{x}_1)) \frac{\partial f_1}{\partial x_1}(h^{-1}(\bar{x}_1), \bar{g}(\bar{x}_1, \bar{x}_2)) f_1(h^{-1}(\bar{x}_1), \bar{g}(\bar{x}_1, \bar{x}_2)) \\ &\quad + h'(h^{-1}(\bar{x}_1)) \frac{\partial f_1}{\partial x_2}(h^{-1}(\bar{x}_1), \bar{g}(\bar{x}_1, \bar{x}_2)) f_2(h^{-1}(\bar{x}_1), \bar{g}(\bar{x}_1, \bar{x}_2)) \\ &\quad + \frac{\sigma^2(h^{-1}(\bar{x}_1), \bar{g}(\bar{x}_1, \bar{x}_2))}{2} h'(h^{-1}(\bar{x}_1)) \frac{\partial^2 f_1}{\partial x_2^2}(h^{-1}(\bar{x}_1), \bar{g}(\bar{x}_1, \bar{x}_2)) \end{aligned}$$

et

$$\bar{\sigma}(\bar{x}_1, \bar{x}_2) = \sigma(h^{-1}(\bar{x}_1), \bar{g}(\bar{x}_1, \bar{x}_2)) h'(h^{-1}(\bar{x}_1)) \frac{\partial f_1}{\partial x_2}(h^{-1}(\bar{x}_1), \bar{g}(\bar{x}_1, \bar{x}_2)).$$

On est dans les conditions de la section précédente. Pour le système (11), on considère le filtre

$$\begin{cases} d\bar{m}_1(t) = \bar{m}_2(t) dt + \sqrt{\frac{2\bar{\sigma}}{\varepsilon}} [dy(t) - \bar{m}_1(t) dt] \\ d\bar{m}_2(t) = \bar{f}(\bar{m}_1(t), \bar{m}_2(t)) dt + \frac{\bar{\sigma}}{\varepsilon} [dy(t) - \bar{m}_1(t) dt], \end{cases} \quad (12)$$

avec condition initiale \bar{M}_0 et où $\bar{\sigma}$ est une constante strictement positive.

Pour le système initial (système (10)), on considère naturellement le filtre

$$\begin{cases} m_1(t) = h^{-1}(\bar{x}_1(t)) \\ m_2(t) = g(\bar{m}_1(t), \bar{m}_2(t)). \end{cases} \quad (13)$$

On a le

Théorème 3.1 Si $\mathbf{E}[\|X_0 - M_0\|^2] \leq C < +\infty$ alors on a les estimations:

$$\mathbf{E}[(x_1(t) - m_1(t))^2] \leq C e^{-ct/\sqrt{\varepsilon}} + C \varepsilon^{3/2}$$

$$\mathbf{E}[(x_2(t) - m_2(t))^2] \leq \frac{C}{\varepsilon} e^{-ct/\sqrt{\varepsilon}} + C \varepsilon^{1/2}.$$

Preuve

Le système transformé (11) vérifie les hypothèses énoncées dans la section précédente et $\mathbf{E}[\|\tilde{X}_0 - \tilde{M}_0\|^2] \leq C$, car la fonction $h'f_1$ est supposée Lipschitzienne. On peut donc appliquer le théorème 2.3:

$$\mathbf{E}[(\bar{x}_1(t) - \bar{m}_1(t))^2] \leq C e^{-ct/\sqrt{\varepsilon}} + C \varepsilon^{3/2}$$

$$\mathbf{E}[(\bar{x}_2(t) - \bar{m}_2(t))^2] \leq \frac{C}{\varepsilon} e^{-ct/\sqrt{\varepsilon}} + C \varepsilon^{1/2}$$

$$\mathbf{E}[(\bar{x}_1(t) - \bar{m}_1(t))^4] \leq C e^{-ct/\sqrt{\varepsilon}} + C \varepsilon^3.$$

D'autre part, puisqu'on suppose que les fonctions h^{-1} et \bar{g} sont Lipschitziennes et h' est bornée, on a les majorations

$$\mathbf{E}[(x_1 - m_1)^2] \leq C \mathbf{E}[(\bar{x}_1 - \bar{m}_1)^2]$$

$$\mathbf{E}[(x_2 - m_2)^2] \leq C \mathbf{E}[(\bar{x}_2 - \bar{m}_2)^2] + C \mathbf{E}[\bar{x}_2^2 (x_1 - m_1)^2].$$

Vu que $\mathbf{E}[\bar{x}_2^4] \leq C < +\infty$, l'inégalité de Hölder permet d'obtenir le théorème. ■

Remarque 3.2 On peut penser à appliquer à ce problème le même genre de démarche que celle utilisée dans la section 2. On obtient alors un résultat analogue au théorème 3.1, mais seulement si on fait une restriction sur la variation des dérivées des fonctions h et f_1 : on exige que les valeurs prises par ces fonctions restent dans un certain voisinage.

Plus précisément, sous les hypothèses

(H8) h est une fonction C^1 , h' est bornée strictement positive et σ est une fonction bornée strictement positive

(H9) f est une fonction C^1 à dérivées bornées et

$$\frac{\partial f_1}{\partial x_2}(x_1, x_2) > 0, \forall (x_1, x_2).$$

Si on note:

$$\phi_X = \frac{\frac{\partial f_1}{\partial x_2}(x_1, x_2)}{h'(x_1)},$$

$$h = \inf_x h'(x), \quad H = \sup_x h'(x),$$

$$f_{12} = \inf_{(x_1, x_2)} \frac{\partial f_1}{\partial x_2}(x_1, x_2), \quad F_{12} = \sup_{(x_1, x_2)} \frac{\partial f_1}{\partial x_2}(x_1, x_2)$$

et on considère le filtre approché

$$\begin{cases} dm_1(t) = f_1(m_1(t), m_2(t)) dt + \sqrt{\frac{2\tilde{\sigma}}{\varepsilon}} \phi_M [dy(t) - h(m_1(t)) dt] \\ dm_2(t) = f_2(m_1(t), m_2(t)) dt + \frac{\tilde{\sigma}}{\varepsilon} [dy(t) - h(m_1(t)) dt], \end{cases}$$

où $\tilde{\sigma}$ est une constante, $\tilde{\sigma} > 0$, on peut démontrer le résultat suivant:

si

$$1 \leq \frac{H F_{12}}{h f_{12}} \leq \frac{(\sqrt{1 + 4(1 + \sqrt{2})} - 1)^2}{4} \quad (1) \quad (14)$$

alors

$$E[(x_1(t) - m_1(t))^{2p}] \leq C_p e^{-c_p \sqrt{\varepsilon} t} + C_p \varepsilon^{3p/2}$$

$$E[(x_2(t) - m_2(t))^{2p}] \leq \frac{C_p}{\varepsilon^p} e^{-c_p \sqrt{\varepsilon} t} + C_p \varepsilon^{p/2}.$$

La méthode qui a permis d'obtenir la condition (14) n'est pas exhaustive. On peut facilement trouver d'autres conditions du même genre pour lesquelles le résultat est encore vrai.

1

$$\frac{(\sqrt{1 + 4(1 + \sqrt{2})} - 1)^2}{4} \approx 1.282$$

4 Commentaires

On a étudié l'erreur de filtrage pour le cas d'un signal de dimension 2 dont une seule des composantes est bruitée et dont seule l'autre composante, faiblement bruitée, est observée. On a obtenu des estimations asymptotiques pour cette erreur qui nous donnent l'ordre de grandeur quand l'intensité du bruit de mesure ε tend vers zero. Une étude asymptotique de l'erreur du filtre approché par rapport au filtre optimal est aussi envisageable (cf. [Picard-89]).

On peut envisager le même type d'étude pour le cas où c'est la composante bruitée du signal qui est observée. Pour ce cas là, dans [YBS], les auteurs proposent un filtre de dimension 2 obtenu formellement d'après un développement asymptotique du filtre optimal.

Références

- [Bensoussan] Bensoussan, A. : *On some approximation techniques in non linear filtering*, Proceedings of June 1986 Workshop, Institute for Math. and Its Applications, University of Minnesota.
- [Fleming-Pardoux] W.H. Fleming, E. Pardoux : *Piecewise monotone filtering with small observation noise*, Siam J. Control, Vol. 27, No. 5, pp 1156-1181, Sep. 1989.
- [KBS] Katsur, R. Bobrovsky, B.Z. and Schuss, Z. : *Asymptotic analysis of the optimal filtering problem for one-dimensional diffusions measured in a low noise channel*, SIAM J. Appl. Math. 44 (1984), Part I: 591-604, Part II: 1176-1191.
- [Picard-86] Picard, J. : *Nonlinear filtering of one-dimensional diffusions in the case of a high signal-to-noise ratio*. 1986 - SIAM J. Appl. Math. (46),6,1098-1125.
- [Picard-88] Picard, J. : *Nonlinear filtering and smoothing with high signal-to-noise ratio*, Stochastic Processes in Physics and Engineering (Bielefeld 1986), Reidel 1988.
- [Picard-89] Picard, J. : *Efficiency of the extended Kalman filter for non linear systems with small noise*, Rapport de Recherche n0. 1068, INRIA, Août 1989.
- [Rogers-Williams] Rogers, L.C.G. & Williams, D. : *Diffusions, Markov processes, and Martingales*, Vol. 2: Itô Calculus, John Wiley & Sons. 1987.
- [Saniuk-Rhodes] Saniuk, J.M. & Rhodes, I. : *A matrix inequality associated with bounds on solutions of algebraic Riccati and Lyapunov equations*, IEEE Trans. Aut. Cont. 32, 8, 739-740, 1987.
- [YBS] Yaesh, I. Bobrovsky, B.Z. & Schuss, Z. : *Asymptotic analysis of the optimal filtering problem for two dimensional diffusions measured in a low noise channel*, SIAM Journal of Appl. Math.. Vol. 50, No. 4, pp. 1134-1155, August 1990.

Bibliographie

- [1] V.E. Benès : *Exact finite dimensional filters for certain diffusions with non-linear drift*, Stochastics, 5, 1981, 65-92.
- [2] A. Bensoussan : *On some approximation techniques in non linear filtering*, Proceedings of June 1986 Workshop, Institute for Math. and Its Applications, University of Minnesota.
- [3] Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, New York, 1968.
- [4] B.Z. Bobrovsky, M. Zakai & O. Zeitouni : *Error bounds for the nonlinear filtering of signals with small diffusion coefficients*, IEEE Trans. on Info. Theory, vol. 34, No 4, July 1988, pp. 710-721.
- [5] M. Chaleyat-Maurel & D. Michel : *Des résultats de non existence de filtre de dimension finie*, Stochastics, 13, 1984, pp. 83-102.
- [6] M.H.A. Davis : *New approach to filtering for nonlinear systems*, IEE Proc., Vol. 128, Pt. D, No. 5, September 1981.
- [7] G. DiMasi & W. Runggaldier, *On measure transformation for combined filtering and parameter estimation in discrete time*, Systems & Control Letters 2, pp 57-62, 1982.
- [8] S.N. Eicher and T.G. Kurtz : *Markov Processes: Characterization and Convergence*, John Wiley & Sons, 1986.
- [9] W.H. Fleming, D. Ji & E. Pardoux : *Piecewise linear filtering with small observation noise*, Proc. 8th Int. Conference on Analysis and Optimization of Systems, Antibes 1988. Lect. Notes in Control and Info. Sci., n° 111, (pp 725-739), Springer, 1988.
- [10] W.H. Fleming, D. Ji, P. Salame & Q. Zhang : *Discrete Time Piecewise Linear Filtering with Small Observation Noise*, Brown University Providence RI 02912, Division of Applied Mathematics, September 1988 LCDS/CCS 88-27.

- [11] W.H. Fleming & E. Pardoux : *Piecewise monotone filtering with small observation noise*, Siam J. Control, Vol. 27, No. 5, pp 1156-1181, Sep. 1989.
- [12] W. H.Fleming and R.W.Rishel : *Deterministic and Stochastic Optimal Control*, Tome 1, Applications of Mathematics, Springer - Verlag, 1975.
- [13] A. Gelb : *Applied Optimal Estimation*, written by the Technical Staff, The Analytical Sciences Corporation, M.I.T. Press, 1974.
- [14] A.H. Jazwinski : *Stochastic Processes and Filtering Theory*, Academic Press, 1970.
- [15] D. Ji, *Nonlinear filtering with small observation noise*, Ph. D. Thesis, Brown University, 1987.
- [16] G. Kallianpur : *Stochastic Filtering Theory*, Applications of Mathematics, Tome 13, Springer-Verlag 1975.
- [17] R. Katzur, B.Z. Bobrovsky & Z. Schuss : *Asymptotic analysis of the optimal filtering problem for one-dimensional diffusions measured in a low noise channel*, SIAM J. Appl. Math. 44 (1984), Part I: 591-604, Part II: 1176-1191.
- [18] F. Le Gland : *Estimation de paramètres dans les processus stochastiques, en observation incomplète - Application à un problème de radio-astronomie*, Thèse présentée à l'Université Paris IX-Dauphine le 26 Novembre 1981 pour obtenir le titre de Docteur-Ingenieur.
- [19] J. Levine : *Finite dimensional filters for a class of nonlinear systems and immersion in a linear system*, SIAM J. Control and Application, vol. 25, No. 6, Nov. 1987, PP. 1430-1439.
- [20] J. Neveu : *Martingales à temps discret*, Masson & Cie Edit., 1972.
- [21] E. Pardoux : *Filtrage non linéaire et équations aux dérivées partielles associées*, Ecole d'été de Calcul des Probabilités de Saint-Flour, 1989.
- [22] J. Picard : *Nonlinear filtering of one-dimensional diffusions in the case of a high signal-to-noise ratio*, 1986 - SIAM J. Appl. Math. (46),6,1098-1125.
- [23] J. Picard : *Asymptotic study of estimation problems with small observation noise*, dans Stochastic Modelling and Filtering (Rome 1984), Lect. Notes in Control and Inf. Sci. 91, Springer, 1987.
- [24] J. Picard : *Filtrage de diffusions vectorielles faiblement bruitées*, Proc. 7th Int. Conf. on Analysis and Optimization of Systems (Antibes 1986), Lect. N. in Control and Information Sciences, 83, Springer 1986.

- [11] W.H. Fleming & E. Pardoux : *Piecewise monotone filtering with small observation noise*, Siam J. Control, Vol. 27, No. 5, pp 1156-1181, Sep. 1989.
- [12] W. H.Fleming and R.W.Rishel : *Deterministic and Stochastic Optimal Control*, Tome 1, Applications of Mathematics, Springer - Verlag, 1975.
- [13] A. Gelb : *Applied Optimal Estimation*, written by the Technical Staff, The Analytical Sciences Corporation, M.I.T. Press, 1974.
- [14] A.H. Jazwinski : *Stochastic Processes and Filtering Theory*, Academic Press, 1970.
- [15] D. Ji, *Nonlinear filtering with small observation noise*, Ph. D. Thesis, Brown University, 1987.
- [16] G. Kallianpur : *Stochastic Filtering Theory*, Applications of Mathematics, Tome 13, Springer-Verlag 1975.
- [17] R. Katzur, B.Z. Bobrovsky & Z. Schuss : *Asymptotic analysis of the optimal filtering problem for one-dimensional diffusions measured in a low noise channel*, SIAM J. Appl. Math. 44 (1984), Part I: 591-604, Part II: 1176-1191.
- [18] F. Le Gland : *Estimation de paramètres dans les processus stochastiques, en observation incomplète - Application à un problème de radio-astronomie*, Thèse présentée à l'Université Paris IX-Dauphine le 26 Novembre 1981 pour obtenir le titre de Docteur-Ingenieur.
- [19] J. Levine : *Finite dimensional filters for a class of nonlinear systems and immersion in a linear system*, SIAM J. Control and Application, vol. 25, No. 6, Nov. 1987, PP. 1430-1439.
- [20] J. Neveu : *Martingales à temps discret*, Masson & Cie Edit., 1972.
- [21] E. Pardoux : *Filtrage non linéaire et équations aux dérivées partielles associées*, Ecole d'été de Calcul des Probabilités de Saint-Flour, 1989.
- [22] J. Picard : *Nonlinear filtering of one-dimensional diffusions in the case of a high signal-to-noise ratio*, 1986 - SIAM J. Appl. Math. (46),6,1098-1125.
- [23] J. Picard : *Asymptotic study of estimation problems with small observation noise*, dans Stochastic Modelling and Filtering (Rome 1984), Lect. Notes in Control and Inf. Sci. 91, Springer, 1987.
- [24] J. Picard : *Filtrage de diffusions vectorielles faiblement bruitées*, Proc. 7th Int. Conf. on Analysis and Optimization of Systems (Antibes 1986), Lect. N. in Control and Information Sciences, 83, Springer 1986.